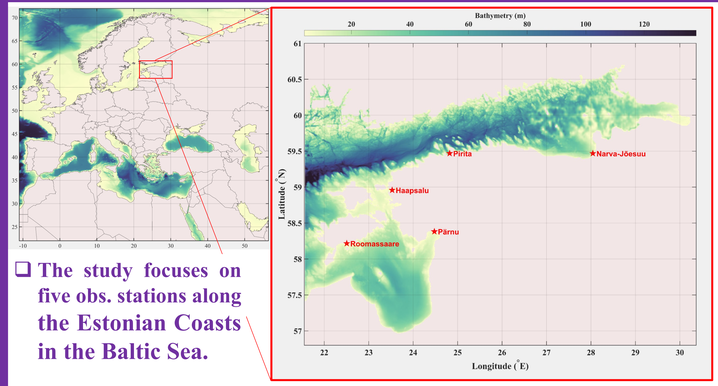
Evaluating the application of deep-learning ensemble sea level and storm surge forecasting in the Baltic Sea

Amirhossein Barzandeh[1], Marko Rus[2], Matjaž Ličer[3,4], Ilja Maljutenko[1], Jüri Elken[1], Priidik Lagemaa[1], and Rivo Uiboupin[1]

- [1]Department of Marine Systems, Tallinn University of Technology, Tallinn, Estonia (amirhossein.barzandeh@taltech.ee)
- [2]Faculty of Computer and Information Science, Visual Cognitive Systems Lab, University of Ljubljana, Ljubljana, Slovenia
- [3]Slovenian Environment Agency, Ljubljana, Slovenia
- [4]National Institute of Biology, Marine Biology Station, Piran, Slovenia



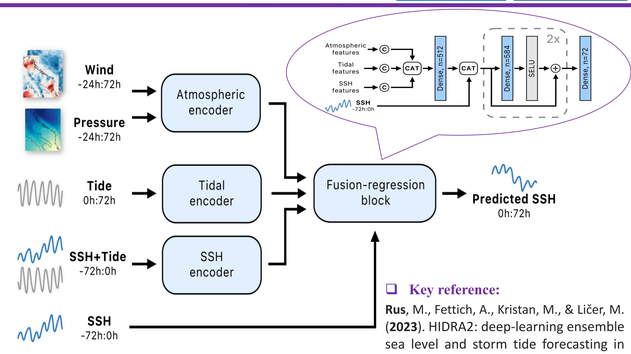
❑ **The study focuses on five obs. stations along the Estonian Coasts in the Baltic Sea.**

❑ HIDRA2 is a second-generation deep-learning model specifically designed to predict sea surface height (SSH). Input data are encoded using three distinct encoders. The resulting outputs undergo recalibration, fusion with past 72-hour SSH data, and regression within the fusion-regression block to generate hourly predictions of SSH for the subsequent 72 hours.

❑ A single forecast run of the HIDRA2 model generates a 72-hour SSH time series for each of the five locations. The computational expenses are minimal, with forecasts for each location taking less than 10 seconds!

❑ In this study, HIDRA2 has been run for one complete year, spanning from April 2023 to March 2024.

❑ For further comparison, HIDRA's performance is assessed against three datasets:
1) outputs from the $NEMO_{EST}$ model run, an operational hydrodynamical model with focused domain on the study area.
2,3) original and detide SSH data from the Baltic Sea Physics Analysis and Forecast, Copernicus Marine Service (CMS), based on NEMO 4.0, which includes T/S profile data assimilation and covers a larger domain encompassing the North Sea and the entire Baltic Sea.

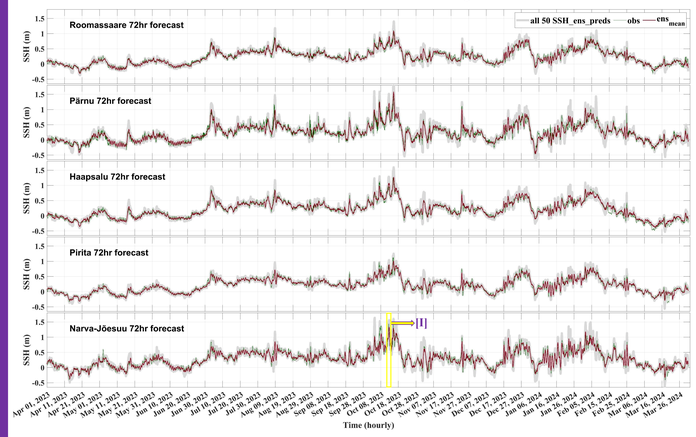


❑ **Key reference:**

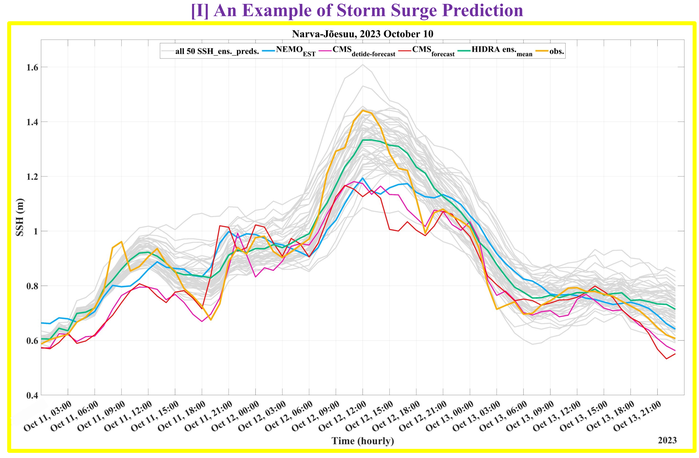
Rus, M., Fettich, A., Kristan, M., & Ličer, M. (**2023**). HIDRA2: deep-learning ensemble sea level and storm tide forecasting in the presence of seiches–the case of the northern Adriatic. *Geoscientific Model Development*, 16(1), 271-288.

https://doi.org/10.5194/gmd-16-271-2023

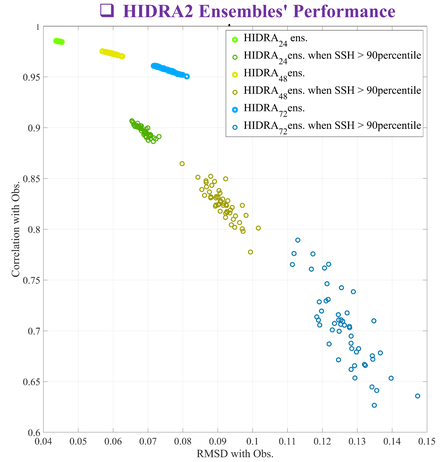
❑ **All observations vs. HIDRA72hr predictions**



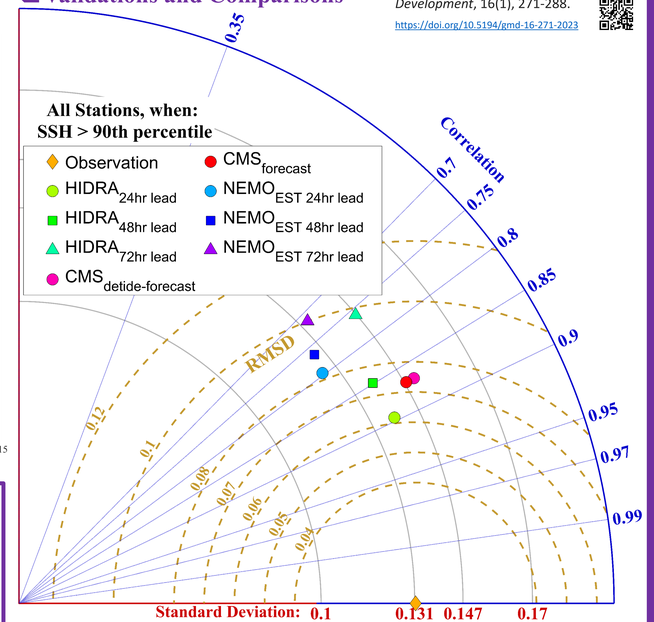
[I] An Example of Storm Surge Prediction



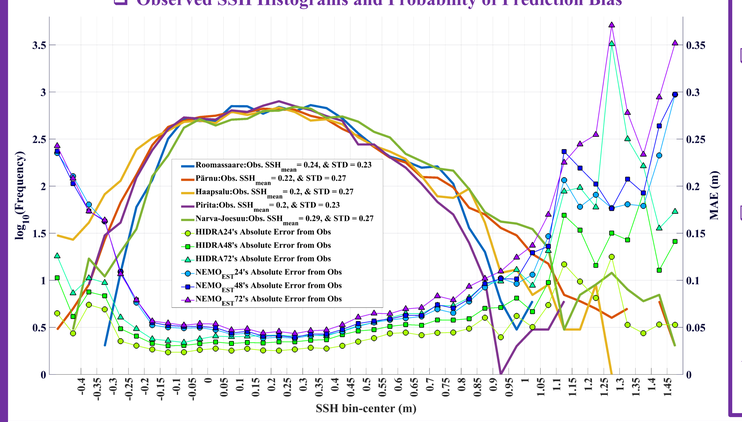
❑ **HIDRA2 Ensembles' Performance**



❑ **Validations and Comparisons**



❑ **Observed SSH Histograms and Probability of Prediction Bias**



In summary, HIDRA2 outperforms other selected models and products in the study area

❑ **HIDRA2's 72-hour forecast, with an averaged RMSD of 0.06 and a correlation coefficient of 0.97, along with a standard deviation closer to observational values, proves to be more accurate than forecasts from $NEMO_{EST}$ and CMS products.**

❑ **HIDRA2's 24-hour forecasts are the most reliable across all the comparisons. HIDRA2's 48-hour forecasts are considered the second most trustworthy and are still superior than the forecasts made by $NEMO_{EST}$ and the CMS products in the study area.**

❑ **In the context of storm surges, comparisons were made based on instances where observed SSH equaled or exceeded the 90th percentile threshold of sorted values within the study period (0.65 m, averaged for all 5 stations in the study area).**

❑ **In storm surges, extending the lead forecast time to 72 hours reveals HIDRA2 to be nearly as accurate as hydrodynamical models. Still, HIDRA2's 24-hour and after that HIDRA2's 48-hour forecasts are the best.**

✓ **This case study holds significance in guiding decision-making processes concerning the integration of deep-learning methodologies into the operational phase of sea level prediction, particularly within the Baltic Sea region.**