EGUsphere, referee comment RC3
https://doi.org/10.5194/egusphere-2022-618-RC3, 2022

Comment on egusphere-2022-618

Anonymous Referee #3

Referee comment on "HIDRA2: deep-learning ensemble sea level and storm tide forecasting in the presence of seiches – the case of the northern Adriatic" by Marko Rus et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-618-RC3, 2022

This manuscript shows an implementation of a neural network (HYDRA2) to predict sea level at the Koper tidal station. This neural network is an improvement over HIDRA1 and it is compared to its predecessor and to a numerical ocean model NEMO. Particularly notable in this manuscript is the detailed validation of the performance of the model. I can recommend publishing this manuscript after minor changes.

To facilitate the reading of the manuscript and the interpretation of the figures and table I would recommend the captions clarify if the authors show an independent validation (data not used during training and not used for the optimization of hyperparameters, if this is the case) or validation with dependent data. Likewise I think it would be useful to mention this also with the skill scores mentioned in the abstract (starting at line 7) whether these error reductions are obtained from the independent test data or not.

I don’t not have any doubts about the scientific soundness of the results, but adding this information would help readers understand the results of the manuscript more quickly.

Minor comments:

Line 5: “single member of ECMWF atmospheric ensemble”: is this the central forecast or any single member (chosen at random)?

Line 47: “HIDRA1 ensemble (Å½ust et al., 2021) is a million times faster than the operational numerical ocean model ensemble based on NEMO engine (Madec, 2016) at Slovenian Environment Agency”: There is not a lot of context to understand this comparison. NEMO will provide you with a sea level estimate over the whole domain. Is this also the case for HIDRA1 or would it provide the sea-level for a single location?

Line 99: “HIDRA2 does not require explicit annotation of whether a location point belongs to land or sea, thus land masks are not generated.”

I am wondering if the land-sea mask would still be a useful feature to provide to the neural network as a wind over land would not generate seiches. I guess that the neural network compensates for this by learning the land-sea mask internally.

Line 103: “three-days prediction lead time” I think that your ML model will give in one application the full 3-day time series. Can you confirm? Or do you rather need to apply the ML model iteratively to obtain the 3-day time series? Can you also clarify this in the manuscript?

104: “full ECMWF three-day forecast” -> “full” refers to the full ensemble (i.e. all ensemble members)?

143: “prototype matching layer” Can you provide more information and a reference ?

section 4.1.1: this is an interesting and surprising result. Can the authors speculate why this is the case? (predicting full SSH leads to better results for extreme events). Could it be that the neural network internally limits its output range when working on anomalies? Do you expect this outcome to remain should one have more training data?

Section 4.1.2 “Ablation study”: can you clarify that you retrained the network for the different test cases (without tides encoder, without atmospheric encoder,..) or you do rather zero-out the output of the corresponding encoder without re-training.

Typo:

Line 205: 1°/24 -> 1/24 °