

Dear editor, dear reviewer.

We thank the reviewer for their encouraging remarks. We have added further explanations to many sections of the manuscript and we would like to make a very short synopsis of what was done during this revision.

1. We have found a minor date-matching error in our preprocessing scripts that had to be corrected. This required a full re-training of all HIDRA2 versions, including ablations and HIDRA1. While the numbers in Tables 1 and 2 are slightly altered due to retraining, the relative performance of HIDRA2 is negligibly altered by this change and all conclusions regarding performance from the original manuscript still hold.
2. During the code review, we noticed that F1, precision, and recall scores are not correctly calculated due to an error at the boundaries of each prediction and we corrected the code. The correction does not change the conclusions regarding performance made in the original manuscript. We have updated and published the code for the calculation of the scores.
3. We have thoroughly rewritten and expanded the section of the ablation study.
4. We have obtained SSH timeseries from the high resolution SCHISM model and repeated all of the analyses with these timeseries. As the reviewer suspected, SCHISM performs better than the forecasting run of NEMO, but HIDRA2 still performs better than both at the tide gauge in Koper.
5. Terminology is now consistent with Gregory et al., as suggested.