

Reviewer 2

The revision is good, and the paper will be ready for publication after minor revision.

I still have a major issue with terminology. The experiments where observations are sampled from the same model are correctly labeled as identical twin experiments. However, the experiments where observations are sampled from HYCOM and assimilated into ROMS are not fraternal twin experiments. The term 'fraternal twin' needs to be dropped in describing these experiments.

The category 'fraternal twin' is reserved for the case where observations are sampled from one model, and then assimilated into the same model that is set up with a substantially different configuration. This is possible with HYCOM, for example, because this model contains multiple choices of numerical schemes and subgrid-scale parameterizations. These different choices can enable the version used to generate observations behave much differently from the version used to assimilate the observations. Even though the same model is used, the two different configurations can be set up to substantially behave like different ocean models.

R: We will follow the suggestion and replace all occurrences of 'fraternal twin' with 'non-identical twin' in the revised manuscript. The following explanatory text will also be added: (new text is underlined here):

"If the chosen "truth" and forecast runs are from same model but with perturbed initial, forcing or boundary conditions, the method is referred to as 'identical twin' approach; if two different model types or significantly different configurations of the same model type (e.g., using different physical parameterizations and/or spatial resolution) are used, the method is referred to as the 'non-identical twin' approach. We note that the approach where the same model type is employed but with sufficiently different configurations is conventionally termed fraternal twin (Halliwell et al., 2014), but here a different model type is used in the 'non-identical twin'."