

## ***Interactive comment on “Evaluation of fraternal versus identical twin approaches for observation impact assessments: An EnKF-based ocean assimilation application for the Gulf of Mexico” by Liuqian Yu et al.***

**Anonymous Referee #1**

Received and published: 10 September 2019

Review of os-2019-85: “Evaluation of fraternal versus identical twin approaches for observation impact assessments: An EnKF-based ocean assimilation application for the Gulf of Mexico”

General comments:

In this manuscript, the authors analyze the differences between the identical and fraternal twin approaches for Observing System Simulation Experiments, using the ROMS model with an Ensemble Kalman Filter (EnKF) in the Gulf of Mexico. They find that the

C1

impact assessment differs in both cases, as the identical twin approach tends to overestimate the error reduction from satellite observations and underestimate the error reduction from vertical profiles of temperature and salinity, compared to the fraternal twin approach.

This manuscript is concise and well written. It provides convincing results to illustrate, for the first time, the differences between identical and fraternal twin experiments using an ocean model. I thus recommend publication, however with minor revision, in order for the authors to better present their results with respect to the reference work by Halliwell et al. (2014, 2015, 2017). Please see my specific comments below.

Specific comments

I find that the approach followed by the authors has a methodological limitation in that, in their fraternal twin approach, they use a data-assimilative model, and that this model has a coarser resolution for the “Truth” than the assimilative simulation. This contradicts the recommendations by Atlas et al. (1997) and Halliwell et al. (2014, 2015, 2017), the work of whom is the reference for the present study. Halliwell et al. (2014), based on Atlas et al. (1997), stated (p. 106): “The established procedures to design and perform OSSEs documented in the atmospheric OSSE literature are summarized by Atlas (1997). The [Nature Run (i.e. the Truth)] is a long unconstrained simulation performed at high resolution using a state-of-the-art general circulation model.” Here, the Truth simulation is not unconstrained, as it is derived from a global operational model that assimilates observations, and it cannot be considered to be at high resolution, since its resolution is  $\sim 9$  km ( $1/12^\circ$ ), larger than the data-assimilative ROMS model resolution (5 km). It is not clear why the authors did not follow the recommendations from Halliwell et al. (2014) and Atlas et al. (1997). They should mention this limitation in their approach in the manuscript, in the methodological section, and include it in their discussion.

Second, I find that the introduction and the discussion sections give a misleading ac-

C2

count of the results and recommendations exposed in Halliwell et al. (2014). In the introduction, at the bottom of p. 4, the authors suggest that Halliwell et al. (2014) recommend investigating the error growth between the various models and observations (l. 61-66), or (“alternatively” l. 67) performing a set of comparable OSSEs and Observing System Experiments (OSEs) using the same data-assimilative model and actual observations. This is misleading, as both steps are recommended by Halliwell et al. (2014).

I find that the account of the recommendations from Halliwell et al. (2014) is more problematic in the discussion. The authors write (l. 355-359): “[Halliwell et al. (2014)] main criterion is that the rate of error growth between simulated and observed states must be similar between the twin framework and reality. However, we found a similar rate of error growth in Sea Surface Height (SSH) in both twin experiments and in reality, yet the identical twin proved problematic. Thus, assessing error growth in just one ocean property appears to have been insufficient.” Not only is the comparison of the error growth one of two main criteria exposed by Halliwell et al. (2014) (see previous paragraph), but Halliwell et al. (2014) never suggested to compare the error budget in only one ocean property, which is what the authors suggest here. Indeed, Halliwell et al. (2014) compared the error budget in SSH, Sea Surface Temperature (SST) and Sea Surface Salinity. The authors’ account of Halliwell et al. (2014) is misleading and they should re-write that part of their discussion to avoid confusion. I also suggest that the authors present the error growth in SST, in addition to SSH, so that their evaluation of the error budget is not performed in one ocean property only. The last sentence of that paragraph (l. 359-362) should also be modified, as Halliwell et al. (2014) recommended a comparison between comparable OSSEs and OSEs as part of the evaluation of the OSSE system, in addition to (and not alternatively to) a comparison of the error rate.

Below are minor specific comments:

- l. 31: Moore et al. (2019) is not in the reference list.

C3

- l. 32-37: It is also possible to keep some of the observations from the pool of observations to be assimilated, to be used for independent assessment of the performance of a data-assimilative simulation. However, this leads to a reduction in the quantity of data that are assimilated. The authors might mention that approach here.

- l. 124-126: Can the authors be more precise about how the model has a tendency to overestimate the Loop Current northward penetration?

- l. 128-145: That part describes the EnKF. Can the authors briefly mention what the specificities of the DEnKF are?

- l. 152-153: Altimetry data are available daily along satellite tracks with a repetitive period of ~10 days for the reference altimetry missions, and the SST data are available daily with higher resolution than  $\frac{1}{4}^\circ$ , in the absence of cloud coverage. The assimilation of weekly maps of SSH and SST at  $\frac{1}{4}^\circ$  resolution is thus a choice of the authors for their experiments, which they should make clear and explain the reason for.

- l. 204: Is the MAD equal to the RMS Error? If yes, I suggest the authors to mention it. If not, I recommend that the authors provide the equation to estimate the MAD.

- l. 248-250: How do the authors explain such a difference in salinity MAD difference in the northeastern shelf of the Gulf, whereas there are no observations assimilated in the area?

- l. 264: Although it is very common, in the scientific literature, to use parentheses to present results from two different datasets or experiments, this way of presenting results is generally confusing and should be avoided, as there is really no reason to use parentheses that way. I recommend the authors to read Robock (2010, <https://eos.org/opinions/parentheses-are-not-for-references-and-clarification-saving-space>).

- l. 324-325: Do the authors have an idea as to why “the additional information content in the subsurface observations within the identical twin system is much smaller than

C4

that for the fraternal twin”? I suggest that the authors discuss this and offer some possible explanation.

- l. 335: Where does this 70 km resolution come from? Is it from the spatial distance between vertical profiles in experiment F3/I3? This should be clarified in the text.

- l. 342-345: Do the authors have an idea as to why the experiment I1 leads to improvement in resolving the small scale processes on the shelf break, in addition to the large scale in the deep Gulf, whereas such improvement on the shelf break was not seen in experiment F1? I suggest that the authors discuss this and offer some possible explanation.

- Supporting Information and Figures: I do not understand why some figures are in a Supporting Information section and others are in the main manuscript. All the figures from the Supporting Information have a comparable role as the figures in the main manuscript. I strongly recommend including all the figures in the main manuscript and get rid of the Supporting Information section.

---

Interactive comment on Ocean Sci. Discuss., <https://doi.org/10.5194/os-2019-85>, 2019.