Referee #1 comments

- In *Italic* : referees' comments
- In normal font : our answers
- In blue: what was added in the text.

---

1. *If I understand well, there is only one simulation with assimilation, for MH. But no simulation with assimilation for PI. Anomalies studied in the paper are between PI no assim, and MH with assimilation. There is a kind of inconsistancy here that you should comment. Unless only anomalies are assimilated, but from line17@p6520, it seems that absolute values are. Could you please justify this comparison between two simulations that does not follow the same experimental protocol ?*

Thank you for noticing this confusing point. Only anomalies are assimilated. This was specified in the original version in section 2.3 (Proxy-based sea-ice reconstruction): "The model-data comparison and the data assimilation are performed using anomalies considering the PI conditions (1850–1900 yr AD) as reference period". Otherwise, we agree that we could not compare the results of the various experiments.

To make it clearier, we have modified the following sentence in section 2.2 (Data assimilation method): After 1 yr of simulation, the likelihood of each particle is computed from the difference between the proxy-based reconstructed and the simulated sea-ice concentration anomalies (6 ka minus PI results), taking into account the errors.

2. *Fig 2 has some features not clearly visible. How many data does not have zero within the error bar ? And for the data that are kept ? As almost all of them does have zero in the error range, are the changes significant ? This point must be addressed.*

In order to improve the figure, in the revised version:

- The size is enlarged.
- The error bars' color is darker to better distinguish which data does have zero within the error bar.

We acknowledge that the sea-ice changes are weak. This makes the quantitative data-comparison difficult and the dataset does not provide a big constraint in the data assimilation process, even if it does influence our model output. To better make the reader aware of these difficulties:

- A sentence has been added at the end of the section 3.1 (Sea-ice changes at 6 ka deduced from observations) : However, out of the 18 proxy-based reconstructions, only 4 (id 2, 4, 5 and 7 on Fig. 1) does not have zero within the error bar, i.e. have a larger signal than their error. This has two implications for the following of this study to keep in mind. First, the potential of this dataset to test the models' performance is weak and second, the constraint applied on the LOVECLIM results during the process of data assimilation will not be large.
- We have added on Fig. 6 the mean signal of each model, which is smaller than the mean signal of the data, itself smaller than the data error. Based on this, we have modified the discussion about the skill of the models at the local scale, saying that it is hard to evaluate the skill of the models quantitatively since the data errors are higher than the signal of both the data and the models. The abstract, the conclusions and a full paragraph in section 3.2 have been updated.

3. *There is no global map of the sea-ice cover for any simulation, and I strongly miss that feature to have of general view of the changes between the different simulations. line25@p6530 (conclusions) states that "The simulated sea ice changes at the MH as compared to the PI period are weaker and*

spatially more homogeneous than the reconstructed ones", and we miss some maps to make up our mind about that. It will also help to relate sea-ice change to atmospheric circulation changes.

This suggestion is followed in the revised version:

- Maps showing the sea-ice edge at the MH and PI as well as the sea-ice concentration anomalies for the annual, winter (March) and summer (September) means have been added for each model. Those maps are available in the supplementary materials, except for LOVECLIM without data assimilation that are in the core of the paper.
- The first 3 paragraphs of section 3.2 (Simulations without data assimilation) were modified in order to take into account these new figures.

4. *In 2.2 Data assimilation method. After reading some of the references, I suggests that you remind the reader that the errors in data are taken into account when eliminating solution too far from data. And give some indication about the method to re-inflate the ensemble. The papers reads about resampling, but I feel that this term is not well chosen to describe the method : perturbation of the atmospheric stream function (?), and not interpolation between particles characteristics.*

The section 2.2 was rewritten taking into account your comments. The term resampling is standard for particle filters (e.g. Dubinkina et al., 2011) and we thus consider that it is well adapted here. This is justified in the revised version of the text. In our case, after re-sampling, the initial conditions of the particles that have just been copied are slightly modified, not by perturbing the atmospheric stream function (as in Mathiot et al. 2013), but by perturbing the sea-surface temperature. This is specified in section 2.2 (Data assimilation method):

LOVECLIM results have been constrained to follow a proxy-based sea-ice reconstruction through a process of assimilation, using a particle filter with re-sampling (van Leeuwen, 2009; Dubinkina et al., 2011), in the same way as in several recent studies (e.g. Goosse et al., 2012; Mathiot et al., 2013; Mairesse et al., 2013). First, an ensemble of 96 simulations (called particles) is initialized from slightly different sea surface temperature for each particle, allowing different time developments. After 1 yr of simulation, the likelihood of each particle is computed from the difference between the proxy-based reconstructed and the simulated sea-ice concentration anomalies (6 ka minus PI results), taking into account the errors. Depending on their likelihood, i.e. their ability to reproduce the signal derived from the available reconstructions, the particles are then either abandoned if their likelihood is low, or kept as a basis for the next year simulation if their likelihood is high enough. In order to maintain a constant number of particles until the end of the simulated period, a resampling, function of the particles likelihood, is conducted annually: the particles with a higher likelihood are copied more times than the others. Finally, the initial conditions of each particles are once more perturbed by adding a small noise to the sea surface temperature of the copies in order to obtain different time developments for the following year, and the whole procedure is repeated sequentially every year until the end of the simulation, here 400 yr.

5. *Are all data points in different model grid points ?*

No, the data id 7 and 8 are in the same grid point, which is also the case of the data id 9 and 10. They are averaged before the evaluation of the likelihood, as the both represent annual mean sea-ice concentration.

References

- Dubinkina, S., Goosse, H., Sallaz-Damaz, Y., Crespin, E., and Crucifix, M.: Testing a particle filter to reconstruct climate changes over past centuries, International Journal of Bifurcation and Chaos, pp. 1–9, 2011.
- Mathiot, P., Goosse, H., Crosta, X., Stenni, B., Braida, M., Renssen, H., Van Meerbeeck, C. J.,

Masson-Delmotte, V., Mairesse, a., and Dubinkina, S.: Using data assimilation to investigate the causes of Southern Hemisphere high latitude cooling from 10 to 8 ka BP, Climate of the Past, 9, 887–901, doi:10.5194/cp-9-570887-2013, 2013.