We thank the reviewer for these useful comments which will certainly improve the manuscript. The reviewers' comments are reproduced in black, and our replies are provided in blue. When necessary, we have also included the modified section of the revised manuscript below our response.

Review RC1 (5 February 2021)

Green et al. presents a new valuable compilation of summer and winter sea ice and SST data for the Southern ocean during the LGM. The paper is, for the most part, quite well written and presented. I very much like that both sea ice and SST data are compiled and then used to evaluate model simulations. The manuscript is however, in its current form, sometimes rather weak on explaining its aims. I do nevertheless find it is suitable in principle for Climate of the Past, after some revision.

Thank you for these comments. In the revised version of the manuscript, we are providing more context and making the manuscript's goals more explicit within the abstract, introduction, and throughout the results.

I have three main comments on this paper.
(1) There is no suitably clear explanation of what is new, or what is being evaluated, or indeed why, in the introduction or abstract. Clearer sign-posting about the paper containing a new compilation of sea ice and SST data for the LGM would also be helpful in the abstract. I found it a difficult to tell whether the new marine data were being evaluated for internal consistency.

We thank the reviewer for this helpful comment. In the new version of the manuscript, we now explicitly state what is new in our analysis. Please refer to our answers below for more detail.

Regarding the internal consistency of the proxy data, the data presented builds on the previous compilation of Gersonde et al. (2005). This compilation was expanded by adding recently published data (Allen et al. (2011), Ferry et al. (2015), Benz et al. (2016), Xiao et al. (2016), Nair et al. (2019), and Ghadi et al. (2020)). This data has been re-evaluated during this work for (1) their chronological framework, (2) their methodology, and (3) their consistency with surrounding data. The studies are based on very similar methodologies ensuring a strong homogeneity.

This was clarified in the Introduction:
"The LGM sea-ice compilation covering the Southern Ocean of Gersonde et al. (2005) is routinely used to provide estimates of LGM sea-ice cover. New Southern Ocean sea-ice data, reconstructed using statistical methods and diatom assemblages, has however been published since then (Allen et al., 2011; Benz et al., 2016; Ferry et al., 2015; Xiao et al., 2016; Ghadi et al., 2020; Nair et al., 2019) thus highlighting the need for an updated compilation."

Or whether CMIP5-PMIP3 models were being evaluated for their ability to simulate Summer and Winter across the Southern Ocean sea ice in the LGM. Or perhaps whether the ability of models to simulate a large (or small) seasonal change in the area or extent of sea ice in the Southern Ocean was being tested. Its also unclear what the motivation is for evaluating models.

We appreciate the reviewer pointing out this confusion. To clarify, our manuscript has several goals: 1) to provide a spatial and seasonal assessment of LGM Southern Ocean sea-ice extent in PMIP3 models (and now also PMIP4 models), 2) to compare available sea-ice and SST records to simulations to assess the skill of each model and of the ensemble, 3) to understand the processes leading to the large range in summer sea-ice extent amongst different model simulations.

This is now clarified in the Introduction:

"Here, we assess the minimum and maximum Southern Ocean sea-ice extent as simulated in LGM PMIP3 and PMIP4 experiments. To better assess intra- versus inter-model variability, LGM sensitivity experiments performed with LOVECLIM are also included. The PMIP3, PMIP4 and LOVECLIM experiments are compared to available sea-ice and SST paleo-proxy data, allowing us to determine the best model-data fit. Combining model and proxy data, we can provide an updated estimate of seasonal Southern Ocean sea-ice cover during the LGM. Furthermore, we analyze the processes that lead to the inter-model differences in summer sea-ice extent at the LGM."

Regarding the motivation for evaluating models, we added a few sentences to the introduction:
"Though paleo-proxy records are an invaluable tool to reconstruct the climate system, they are sometimes scarce or completely absent over entire regions. Climate models can help fill these gaps, as they provide a full 3-dimensional and dynamically consistent representation of the climate system. However, as climate models are not perfect representations of Earth's climate, it is important to continually evaluate their performance."

(2) Closely linked to (1), the choice of models being evaluated is difficult to defend. Evaluating CMIP5-PMIP3 versus CMIP6-PMIP4 – and whether models have improved in their ability to simulation SO sea ice would seem valuable and more easy to understand as motivation. However the paper in its current form does not do this. Instead, it shows results from the older PMIP3 simulations, alongside some more recent LoveClim simulations. At least one of the LoveClim simulations also has a very strange seasonal
cycle of sea ice. The motivation for doing this approach, as opposed to PMIP3 versus PMIP4 is never explained. I would encourage the authors to consider focussing on PMIP4 versus PMIP3/2. More than 10 PMIP4 LGM simulations are available to the authors.

We understand and agree with the reviewer's point here. When we originally started the project (late Fall of 2018), the PMIP4 data was not yet available (it became available late 2020 / early 2021) which is why we decided to use PMIP3 data. However, in the new draft of this manuscript, we have included the available PMIP4 data and compared that against the PMIP3 outputs. Additionally, the strange seasonal cycle in LOVECLIM1 (now referred to as weakNA) was a mistake. In the new draft, we are analyzing the LOVECLIM1 output with LOVECLIM1's summer months now correctly adjusted to February and March.

(3) The section on wind stress results is not particularly helpful. It needs more on how glacial-interglacial wind changes depend on model biases, and the sea ice itself. It can help to start with considered wind velocities above the surface, before then considering wind stress/curl changes. Otherwise results can tend to confuse.

We thank the reviewer for this comment. Our main goal in this section of our study is to understand the dynamic processes that lead to the simulated sea ice extent within each model. We want to understand how sea ice extent is related to wind stress within each of the models; i.e., in a numerical and theoretical world that is not necessarily realistic, but logical within itself. We are therefore not concerned about the ability of the models to accurately represent the winds, we want to know how the winds impact the sea-ice cover within the model framework. We agree that a thorough model-data comparison and evaluation of surface winds would be very interesting, but this is out of scope for this study. Please also refer to the minor comments section, where we provide a more detailed answer to this comment.

Minor comments

L90-_95. Be specific about what 'common' (e.g. 1 in 2 years?) and 'episodic' (e.g. 1 in 10 years, 1 in 50 years?) mean.

It is worth noting that geological data are generally extracted from 1 cm thick slices of sediment that average tens to hundreds of years of accumulation depending on the sedimentation rates. Therefore, though we understand the concern of the reviewer, the % of FCC or diatom transfer function data cannot provide quantitative recurrences.

We accordingly updated section 2.2 of the methods:
"Relative abundances of the indicator diatoms above 3% are thought to indicate the presence of sea ice over the core site (mean sea-ice extent north of the core site) while relative abundances between 1 and 3% suggest the episodic presence of sea ice over the core site (mean sea-ice edge south of the core site but maximum sea-ice edge north of the core site)".

Table 2 – It is most strange that LOVECLIM1 has a SIE minimum in Jan-Feb. Suggest excluding this model from the analysis. Hard to see how the results can be meaningful.

We apologize as there was a mistake in table 2. The SIE minimum in LOVECLIM1 is indeed Feb-Mar, similar to 8 of the other 9 models. This has been corrected in the revised manuscript.

L110 "The proxy"

This has been corrected.

Figure 1 – add latitude grid lines.

We have added latitude grid lines to Figure 1.

L160 onward. Given the freezing point of seawater is _-1.8C, I find the lack of <0C datapoints a little strange.

First, the paleo SST data is based on several approaches, including diatom transfer functions and radiolarian transfer functions (Gersonde et al., 2005). Micro-organisms develop during the sunlit period (spring to fall), when sea ice retreated and SSTs are back above the freezing point. Indeed, SST as low as -1.8° C would indicate yearly round sea-ice coverage during the whole LGM at that location. Such a situation would prevent any plankton production. This is why summer SST is the first explanatory variable for the distribution of many diatom species (Esper et al., 2014).

Second, modern diatom databases cover a SST range from -2°C to 20°C, winter sea-ice concentration from 100% to 0% and summer sea ice concentration from 40% to 0% (Armand et al., 2005; Crosta et al., 2005; Esper et al., 2010, 2014). It is probable that transfer functions would reconstruct < 0°C LGM SST, and high sea-ice concentrations, if adequate cores were available. For example, a diatom transfer function estimated LGM SST as low as -1°C at 60°S-5°W off the Weddell Sea (Gersonde et al., 2003). This core was not used in Gersonde et al. (2005), and therefore in the present compilation, because of its weak chronological constraint.

Would be useful to have some discussion of the relationship between the freezing point (_-1.8C) and the model and obs datapoints here.

We discuss this relationship more within section 3.3 of the results, shown below:

Results Section 3.3:

"Due to biological limitations, diatom transfer functions are mostly available in regions with low sea-ice cover. As such, our proxy summer SST compilation only contains two locations with SST temperatures below 0˚ C. With limited proxy SST data near the freezing point, we instead assess the model-data fit at the 1˚ isoline."

Figure 2. Consider adding a -1.8C or -2C line on panel (c).

We included a -2˚ C line in the revised Figure 2.

Figure 2. Obs data is too faint on panel (c). Make this more visible please.

This is noted and has been fixed.

Table 3 – order of columns seems strange. Maybe should be Winter: SI edge, SIE, SIA, Obs agreement, SST, SST agreement; then Summer: SI edge, SIE, SIA, Obs agreement, SST, SST agreement.

We agree with this suggestion. The table has been amended as suggested.

Figure 3 – Cannot see Southern Ocean temperature variation with current colorbar. Change scale to -2 to +8? Figure not useful with this colorscale.

We have fixed this by using a non-linear color bar.

Line 229-253 There are three main problems with this section.
Firstly it currently lacks adequate discussion of the fact that SHWs are too poorly simulated by most PMIP3 models to be able to straightforwardly interpret wind results.

We want to note that our main goal here is to analyse the internal dynamic processes within each model framework, to better understand summer sea ice extent, and what is driving this extent, within each simulation. Hence we were not particularly concerned about the ability of the models to accurately represent the winds. We have made this goal clearer but also highlighted the fact that there are significant uncertainties in both reconstructions and model simulations. The updated sections of our results and discussion are below.

Results section 3.4:
"While latitudinal position and magnitude of southern hemispheric westerlies at the LGM is poorly constrained (Kohfeld et al., 2013; Sime et al., 2016), here we want to assess the impact of the simulated windstress curl on ocean dynamics in each model. We thus use the simulated windstress outputs to estimate the location and strength of the SO upwelling, and its potential impact on sea-ice cover. "

Discussion:
"Given the uncertainties that surround the magnitude and the position of the Southern Hemisphere westerlies at the LGM (e.g., Kohfeld et al., 2013; Sime et al., 2016), this casts additional uncertainties on the location of the austral summer sea-ice edge."

Secondly, the discussion needs to take account of the control that sea ice exerts on SH winds.

We now also mention that the wind-sea ice relationship goes both ways, and the presence or absence of sea ice can also influence winds. We however want to reiterate that we are more interested in the impact of Southern Ocean upwelling (driven by SH winds) on sea ice. This is the reason we are looking at the windstress curl. The updated sentence is within section 3.3 of the results and included below in response to the next comment.

Thirdly, by focusing on stress and curl, rather than simpler above surface speed/velocity measures (which are less affected by surface roughness) it leads to a discussion that probably confuses the reader more than clarifying anything. Suggest re-reading relevant SO/sea ice wind papers by Bracegirdle et al (2013, 2018 and others), Sime et al (2013, 2016), and Kidston et al (2011, The influence of Southern Hemisphere sea-ice extent on the latitude of the mid-latitude jet stream, Geophys.
Res. Lett., 38, L15804, doi:10.1029/2011GL048056). Rewrite or remove this section?

We understand the point of the reviewer and agree that the cited studies highlight the impact of changes in Southern Ocean sea-ice extent on the position and strength of the SH westerlies. However, here we are interested in the impact of Southern Ocean upwelling, and associated divergence of surface waters, on sea ice. We have modified the caption of new figure 4 as well as the associated text as follows:

Results section 3.3:
"The strength and location of the southern hemispheric westerly and polar easterly winds impact Southern Ocean circulation, sea ice transport and therefore sea-ice distribution (Purich et al., 2016; Holland and Kwok, 2012). On the other hand, the presence or absence of sea ice also has a direct influence on surface winds (Kidston et al. 2011; Sime et al. 2016). The divergence created by the wind stress curl over the Southern Ocean leads to an upwelling of warmer deep waters and thus heat loss to the atmosphere. This upwelling can therefore also impact Southern Ocean sea-ice distribution."