Author Reply to the Review Comments by **Lenneke Jong (Referee #1)**

on the manuscript

How precipitation intermittency sets an optimal sampling distance for temperature reconstructions from Antarctic ice cores

by Thomas Münch, Martin Werner, and Thomas Laepple,
submitted to *Climate of the Past* (https://doi.org/10.5194/cp-2020-128).

Thank you very much, Lenneke Jong, for the time you spent on reading and reviewing our manuscript. Below we include a point-by point response to both the general and all specific comments. The original referee comments are set in normal black font, our replies in blue, and suggested changes to the manuscript are shown by citing the manuscript text on a gray background with changes in red.

General Comments:

This manuscript investigates the possibility of selecting an optimal set of ice core locations to best reconstruct the temperature record at a target Antarctic site. They use an isotope enabled climate model to derive what they term a "sampling correlation structure" of sampling locations from concentric rings radiating out from a target site. Their findings that the optimal reconstruction can be obtained by combining a local core with another 500-1000km away to decrease the noise in the records due to precipitation intermittency is particularly interesting.

Thank you; this indeed might be the most striking result.

The paper itself is quite well structured, however, I think it would benefit from including the details of the conceptual model earlier into the main text of the paper when the decorrelation length scales relating to the sources of noise are introduced. A diagram would also be of great help. It is probably beyond the scope of this paper but I would have liked the authors to comment on the relative importance of the different sources of noise in the isotopic signal and which ones are dominant depending on location in Antarctica (eg coastal vs inland).

We do not fully agree. We think the reviewer refers to the introduction where we discuss the different noise sources (P2 LL27–45). However, firstly we argue here already that each noise source should exhibit a characteristic spatial scale of influence or decorrelation length. Secondly, while we think that, based on the overall review comments, the conceptual model does need some more introductory motivation, we do not think that the introduction is the right place for mentioning already more details of the model. We also think that the methods section is neither appropriate, since the conceptual model is used in order to interpret our results, and not as a method to produce the main results. We therefore suggest that we give more space to motivating the conceptual model and our assumption of radial symmetry at the beginning of section 4.1. Additionally, we will add the following figure to the appendix which graphically illustrates the different decorrelation lengths and processes in the conceptual model:

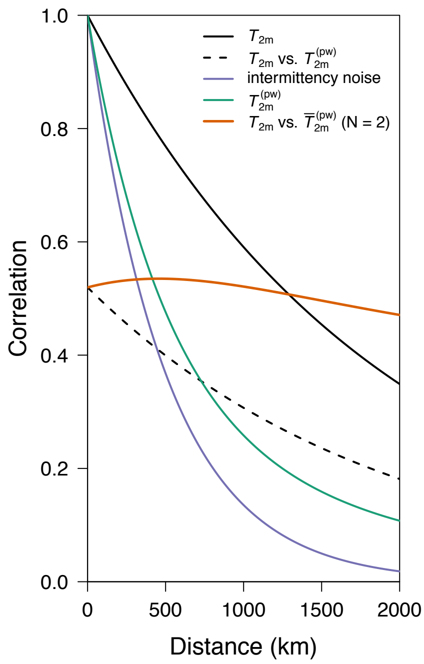


Fig. 1. Illustration of the decorrelation lengths in the conceptual model. Shown are as a function of distance the correlation between two temperature time series (black), between a temperature and a precipitation-weighted temperature time series (dashed black), between the intermittency noise (violet), between two precipitation-weighted temperature time series (green), and between a target temperature time series and the average of two precipitation-weighted temperature time series (orange), one located at the target site and one located a certain distance away from the target site as indicated on the x axis. Parameters are taken from the DML region.

Regarding the relative importance of the different noise sources: Previous studies have shown for the EDML site in East Antarctica that stratigraphic noise amounts to approximately up to 50 % of the total isotope variance at the seasonal time scale (Münch et al., 2016), however, quantitative estimates for other Antarctic regions are still missing. A similarly high relative contribution is expected from precipitation intermittency (Laepple et al., 2018), which probably has a larger impact further inland than compared to coastal regions (Casado et al., 2020). We will add a short discussion of these results to the introduction.

I found some of the discussions of study regions and target regions quite confusing and could do with some clearer explanations. I also thought a motivating example with actual ice core data at one of the target sites to demonstrate the reduction in noise would be very helpful, and would be useful in quantifying how much improvement in the temperature reconstruction is obtained with additional cores when all the other confounding factors are included in the isotope signal. They state that including these sources of noise, eg isotope diffusion is outside the scope of this study, but and I would like to see how the results hold up when the real data is included.

We are sorry that some of our explanations regarding study regions and target sites were not clear enough and we will improve the text as illustrated in our answers to the specific comments.

We agree that testing our results on real ice core data would be an ultimate goal. Such a test would include to find appropriate ice core data and suitable instrumental temperature records, which are sparse on Antarctica. We thus think that this is clearly a study on its own and much beyond the scope of this manuscript, also given the manuscript's current length and number of figures.

The authors are clear in the assumptions going in to the conceptual model and the analysis overall. They are also clear on the limitations of trying to use this optimal sampling correlation structure in the real world.

Overall, the manuscript is written clearly, and is of good quality and is of scientific interest, especially to the ice coring an palaeoclimate community. I have listed a few specific comments below that I believe would help improve the readability of the paper and recommend publication once these issues are addressed.

Thank you. We are happy about this positive evaluation and are confident that addressing the specific comments will further improve the manuscript.

Specific Comments:

Lines 13-14: Remove final sentence as I didn't see the application of this technique to anything other than temperature reconstructions from ice cores discussed.

You are correct that in the paper our technique for assessing the optimal sampling strategy is only applied to temperature reconstructions from ice cores. Therefore, this final sentence of the abstract was meant to be an outlook to emphasise that our technique is however general enough to be extended to other palaeoclimate (temperature) proxies that face similar problems, i.e. noise sources working on different spatial scales. In addition, we do pick up this topic in the final part of the Conclusions (LL 353–357).

Unless the editor explicitly disagrees, we would keep this sentence/part in the manuscript as a reference for the broader palaeoclimate community, but we suggest to rephrase the mentioned sentence to more clearly articulate that it is being meant as an outlook for further studies:

[...] It also broadens our knowledge on the processes that shape the isotopic record and their typical correlation scales. Finally, many palaeoclimate reconstruction efforts face the similar challenge of spatially correlated noise, and our presented method could directly assist further studies in determining optimal sampling strategies also for these problems.

And in the Conclusions (L355 et seq.):

This likely applies to various marine as well as terrestrial proxy types, and our strategy thus might offer a step forward in the best use of sampling and measurement capacity for quantitative climate reconstructions, which needs to be investigated in further studies.

section 2.3.3: Some motivation and justification for the sampling scheme of rings and selection would be helpful. It was also not clear at times if N was referring to the number of grid cells, or rings and also is presumably different to N_{grid} mentioned later on line 132 onwards.

We chose the sampling scheme of rings since it provides a computationally efficient way to estimate a statistically solid average correlation as a function of distance between the averaged locations; see also our more detailed answer on the similar issue raised by the second reviewer.

In general, in the manuscript the symbol N (with or without subscript) always refers to a particular number of model grid cells, but we agree that our use of this terminology may have been unclear at times, since there are different contexts in which we use this symbol. To make it clearer for the reader, we suggest to adjust the terminology as follows:

- In section 2.1 (LL82–84), the number $N_{grid} = 442$ refers to the total number of grid cells which lie, in our specific model simulation, on the Antarctic continent and which are used for all our analyses (since, obviously, you cannot drill ice cores on marine sites). Since this number is not referred to at any later stage in the manuscript, we suggest to drop the term N_{grid} here in L84:
  [...] extracted from the total number of 442 model grid cells that are available for the Antarctic continent (Münch and Werner, 2020).
- In section 2.3.4 (LL129–139), the number N_{grid} refers to the number of grid cells which lie within our defined DML and Vostok study regions and which are used as temperature target sites. To improve clarity, we will, similar to above, refer to the number of grid cells within the two study regions explicitly but without using the symbol N_{grid}, and will rephrase the respective text passages as follows:
  (L132) We define the DML region as the area of ±17.5° longitude and ±5° latitude around the European Project for Ice Coring in Antarctica (EPICA) DML site (EDML; 75° S, 0° E; Fig. 1), consisting of 26 model grid cells.
  (L135) For the Vostok region, we choose an identical latitudinal and longitudinal coverage ~~(N_{grid} =~~

~~30)~~ with respect to the Vostok station (78.47° S, 106.83° E; Fig. 1), covering 30 model grid cells and encompassing the sites of the deep Vostok and Dome C ice cores and of several shallower cores (Stenni et al., 2017), [...].

- Throughout the manuscript, the term *N* (i.e. without subscript) is used to refer to the number of grid cells (i.e. time series at the locations of these grid cells) which are averaged and correlated with a target site temperature time series. To improve clarity, we will replace *N* by N_ℓ (ℓ for "locations") and introduce this terminology explicitly in section 2.3.1.

section 3.1 I would have liked to see what level of significance is attached to each correlation coefficient reported.

The correlation values are all highly significant; please see our answer below to your remark on Fig. 3.

Fig 2: (And other figures) Use of diverging colour map for only positively increasing correlation or lengths scales is a bit distracting. Consider only using reds for instance.

We agree and will use a sequential color scheme (red hue) for Fig. 2 and for the correlation map figures (Figs. 3, 4, 6, A1, and A2).

Fig 3. This is an important figure in that it shows the precipitation weighting being important in the correlation. I would like to see some indication of where these correlations are significant to (eg $p<0.05$). It would be nice to see another map explicitly showing the difference in correlation coefficients between the two, as it looks to be some regions where there is no difference at all.

We will add the map showing the differences in correlation (Fig. 2 below) to the manuscript as part of manuscript Fig. 3. This map nicely illustrates that the correlations tend to remain unaffected mostly in the coastal regions of Antarctica, where precipitation intermittency is expected to be less important (Casado et al., 2020).

However, indicating the significance of the correlation coefficients in the maps does not make sense statistically, since the correlation values are all highly significant ($p <<< 0.01$) given the long time series (1200 data points). Even if one accounted for autocorrelation of the data, the significance should still remain very high across all grid cells.

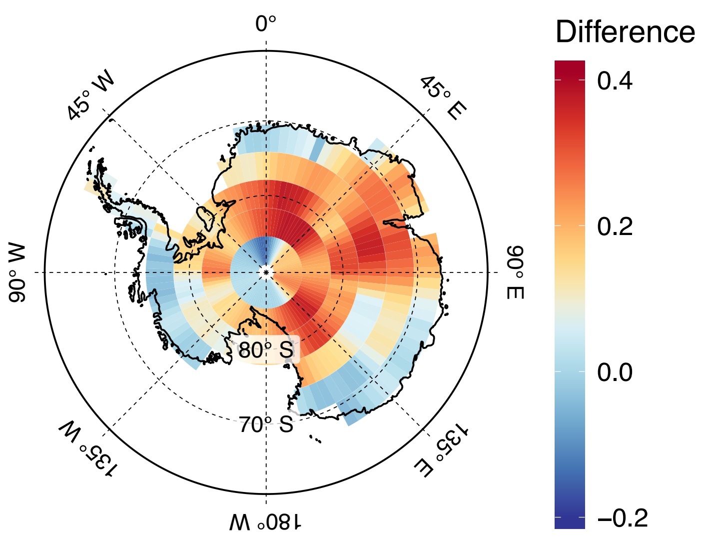


Figure 2. The difference in correlation coefficient between using unweighted and precipitation-weighted 2 m temperature time series for the correlation with the local 2 m temperature (i.e. the difference between manuscript figure 3b and 3a).

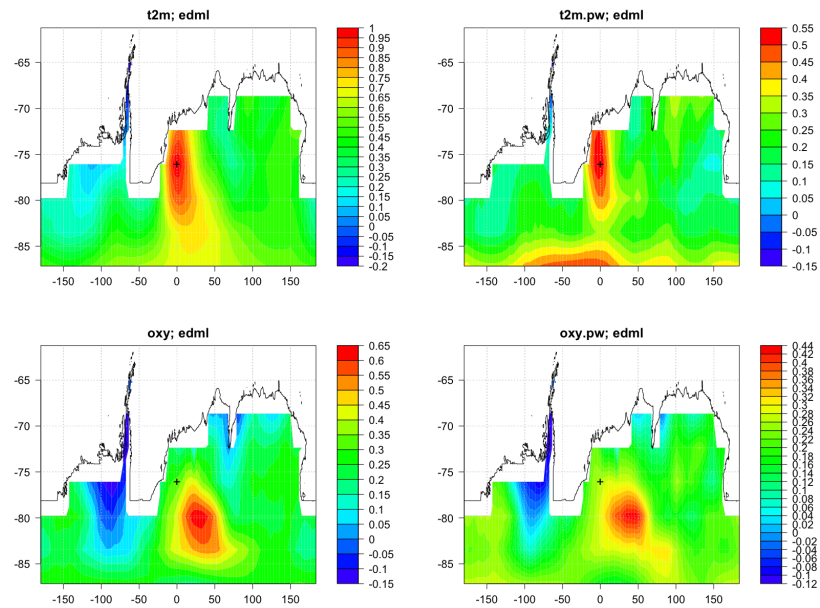
3.3 Optimal ice-core sampling structures line 187 "we compute the mean of correlation results obtained between a target site temperature and individual grid cells in order to reduce local variability in the model data" Is this what you really mean, averaging the correlations? This seems like a strange thing to do if you are looking to maximise the correlation overall?

We agree that our explanation might be be ambiguous or misleading here. Of course, we first average the *N* isotope time series, taken from *N* grid cells, and then compute the correlation of this averaged record with the target site temperature time series. We iterate this approach over all possible combinations (or over a finite number of Monte Carlo combinations) of drawing *N* grid cells from a given ring combination, and only then average the correlations from these iterations to obtain the mean correlation for this ring combination. We then analyse the next ring combination. This approach is in more detail explained in Sect. 2.3.4. We will revise the text in Sect. 3.3. to better and unambiguously summarise this approach. In addition, we will take care that the methods text in Sect. 2.3.4 is also clearly understandable.

I would also like more comment on how much of the regional difference is lost by averaging to only get a function of radial distance. As the DML and Vostok results differ and you suggest there are regional differences. Surely the correlations are not radially uniform?

We agree that the correlations do not need to be radially uniform. From physical arguments we expect, however, that the first-order spatial correlation patterns are largely invariant against rotation. This is indeed the case for the temperature field, as the correlation maps show for the EDML and Vostok sites (Fig. 3 below), and similar patterns are observed for other Antarctic regions. However, for $\delta^{18}O$, and also partly through the effect of precipitation weighting, indeed rather strong radial asymmetry can occur. Nevertheless, still a contribution from radially symmetric patterns may exist, and our approach of a radial averaging is based on assuming such symmetric contributions. This can be motivated by the fact that in real world applications one may not necessarily know in which direction the correlation pattern is maximal, so that radial symmetry is the most straightforward assumption. And indeed our results suggest that a gain in correlation can be achieved nevertheless, when we use optimal core locations based on an analysis assuming radial symmetry, despite the actual form of the spatial correlation patterns.

a.



b.

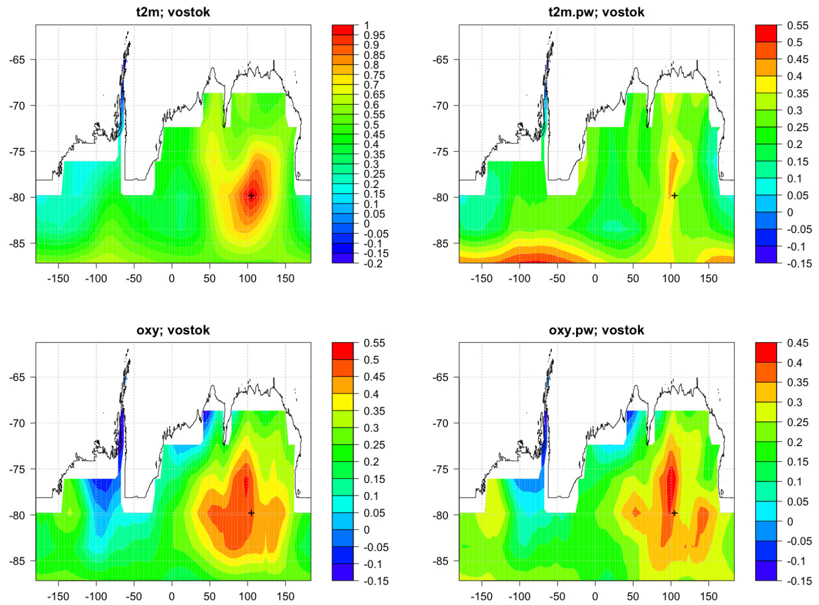


Figure 3. Smoothed maps of the correlation between the temperature time series at the target sites EDML (a) and Vostok (b) and different climate model fields: temperature, precipitation-weighted temperature, $\delta^{18}O$ and precipitation-weighted $\delta^{18}O$ (from upper left to lower right panel, respectively). Note that the correlation patterns for temperature (upper left panels in a and b) are actually nearly radially symmetric around the target site and only appear elongated in North–South direction due to the map projection. For inclusion of this figure into the manuscript we would use a proper colour scheme and a polar projection, similar to manuscript Fig. 4.

At this point, we think it would be most beneficial for the reader to include the presentation of Fig. 3 in the manuscript (e.g. in a new subsection after 3.1) and to use it as a starting point to motivate the ring sampling in section 3.3 as well as for the discussion (Sect. 4.1). However, this would of course extend the manuscript substantially and it would also increase the number of figures, and we thus would leave it to the editor to decide whether this extension is reasonable.

Fig 4. It would be also useful to have the concentric circles marked around the target sites. Can you also comment on why the locations shift when more cores locations are added? That is, is the location in the N=1 case also included in the N=3 case as it looks as though they have moved slightly. These segments don't seem to correspond to the regions mapped out by the black polygons in Fig. 1 so it is not clear to me what the study regions actually are. Does it mean that the cores in the N>1 cases can be from outside of those black polygons too as appears to be the case here?

We use the study regions (the black polygons in Fig. 1) only to define regions from which we select temperature target sites, with respect to which we conduct our analyses and across which the results are then averaged to obtain regional estimates, such as for Figs. 5, 6, and A1. Here, for Fig. 4, we use only a single temperature target site from within the study regions, namely the EDML site (Fig. 4a–c) and the Vostok site (Fig. 4d–f). The $\delta^{18}O^{(pw)}$ model grid cells that we average and correlate with the target temperature time series can, however, indeed be selected from a wider region than the study region, namely from the 2000-km circles around the target sites (as explained in Sect. 2.3.2).

We will add the line of the 2000-km circles around the target site to the plots and explain in the figure caption that the $\delta^{18}O^{(pw)}$ grid cells ("ice cores") can be chosen from within these circles.

Yes, you correctly observe that the optimal location (model grid cell) in the N=1 case for EDML is no longer included in the N=3 or N=5 case, while this is the case for Vostok. However, this is simply by chance: while for N=1 it is computationally easy and fast to find the best correlating grid cell within the 2000-km circles and so panels (a) and (d) display the "true" optimal location, for N=3 and N=5 we need to randomly select and average grid cells, using 10^5 iterations. The displayed locations is the best configuration of these iterations, but does not necessarily need to be the "true" optimal configuration. However, in terms of correlation value with the target site temperature, the value from the best iteration should be very close to the correlation value for the true optimal configuration due to the large number of performed iterations.

Fig 5. The black dashed line indicating the exponential fit is not included in the figure legend, only the caption. The dots in the plot don't appear to be at the expected 0, 250, 500km marks, but are a bit offset, is this deliberate? Again, I question the averaging step around the whole 250km rings, but it would be nice to see the spread in the correlation as a function of distance too. I am confused at what is meant by "all respective target sites in the DML and Vostok region", how are there more than one target sites for each region, are these different that the two crosses shown in Fig. 4?

We will add the explanation for the dashed line to the figure legend and include the correlation scatter (standard deviation of the correlation results across the different target sites within the study regions); see the revised plot in Fig. 4 below. Yes, the dots in the plots are deliberately placed in the middle of the ring bin borders, i.e. at distances of 125, 375, 625, ... km, since each correlation value is an average value across a ring bin (from 0–250, 250–500, 500–750, ... km) by averaging the individual correlations obtained between the target site in the centre and all grid cells that lie within each bin.

Regarding the target sites your comment shows that there is a clear need for us to better explain this concept in the revised manuscript. We use "target site" as the term to denote a model grid cell from which we use the temperature time series (T_{2m}) to correlate all other grid cells and variables with, and which defines the centre of the ring bins. For a single target site we can study the average correlation with distance similar as shown in Fig. 5. But Fig. 5 involves a second averaging step: To improve statistics, and to obtain regional estimates, we use all model grid cells within our defined DML and Vostok regions as target sites, one after the other, to get the correlation–distance dependencies for each target site and for each variable (T_{2m}, δ^{18}O, $\delta^{18}O^{(pw)}$). Then we average all these 26 (DML region) and 30 (Vostok region) curves that we obtain for each variable, respectively, to produce the curves shown in Fig. 5. The same approach is used for the results shown in Figs. 6 and A1.

We will expand the explanations in sections 2.3.1 and 2.3.4 to clarify the concept of "target sites" and of "the averaging across target sites" to the reader.

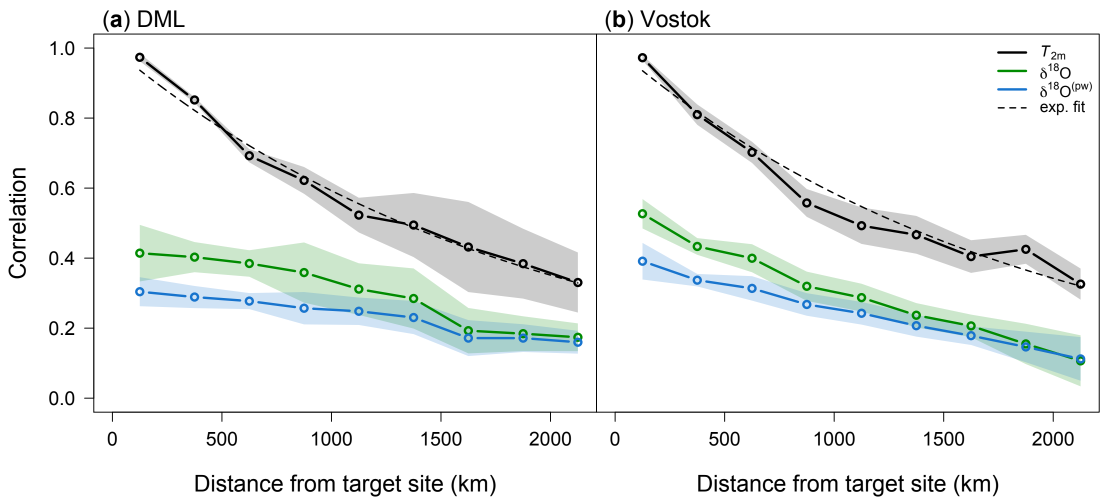


Figure 4. Revised version of manuscript Fig. 5 with an estimated regional scatter of the correlation coefficient included.

Fig 6. These are interesting and show the clear difference for the precipitation weighted δ^{18}O and T_{2m} for DML, but why is the Vostok case relegated to the appendix? The fact that they are different is an interesting result.

We fully agree that the difference between the DML and Vostok regions is an interesting result. When writing up the manuscript, we were concerned that we could overload the results section with too many coloured contour plots, which is why we moved the Vostok plots to the appendix. However, we are happy to combine both regions into a single Figure 6 (similar to Fig. B1) and remove Appendix A.

Fig 7. The way this figure is arranged, is the top row, marked rank 5, that which has the max correlation, or is the rank 1 row the highest correlation row? I find it very curious that in the Vostok region the optimal arrangement comes with no local sites in many of the cases

The maximum correlation correponds to the combination/row that is marked as rank 1; we will add a clarifying remark to the figure caption.

Yes, it is indeed a curious result that in the Vostok region the optimal arrangement comes mostly without local sites, but we see no evidence for not trusting this result in view of the agreement between N = 3 and N = 5 and given the statistics from the large number of sampled locations in our ring sampling scheme.

line 220: Suggests that regional differences play a part here and I would like a comment on what those differences are (elevation, distance from coast etc?)

We can only speculate what possible reasons could there be for the difference in temperature–isotope correlation between the DML and Vostok regions (see Fig. 5 in the manuscript). One factor might indeed be the larger elevation and distance from the coast, i.e. a stronger continentality at Vostok, and related to this, differences in the distillation paths of the transported vapour.

Line 232: Does the averaging have an effect on the significance of the correlations?

As pointed out above, there is statistically no sense in studying the significance of the correlation coefficients given that each time series has such a large number of data points.

Fig. 8 (b) I assume the colours on the histogram are the same as in (a), but please add the legend anyway. Can you comment on the low correlation outliers for the EDML case.

Yes, the colours are the same in (b) as in panel (a); nevertheless, we will add a second legend to panel (b). The low correlation outliers in the EDML case stem from the grid cell combinations which include one anomalous grid cell located at ~ 72.4 °S, 22.5 °E; see Fig. 3 and our reply to a respective comment by reviewer #2. We will have a deeper look into this issue.

Section 4.1 I found most of the discussion very clear and thorough, but re-iterate that a good schematic diagram to illustrate the length scales in the conceptual model would be very useful.

Thank you; please see our reply to your General Comments and Fig. 1 there.

References:

Casado, M., Münch, T., and Laepple, T.: Climatic information archived in ice cores: impact of intermittency and diffusion on the recorded isotopic signal in Antarctica, Clim. Past, 16, 1581–1598, https://doi.org/10.5194/cp-16-1581-2020, 2020.

Laepple, T., Münch, T., Casado, M., Hoerhold, M., Landais, A., and Kipfstuhl, S.: On the similarity and apparent cycles of isotopic variations in East Antarctic snow pits, The Cryosphere, 12, 169–187, https://doi.org/10.5194/tc-12-169-2018, 2018.

Münch, T., Kipfstuhl, S., Freitag, J., Meyer, H., and Laepple, T.: Regional climate signal vs. local noise: a two-dimensional view of water isotopes in Antarctic firn at Kohnen Station, Dronning Maud Land, Clim. Past, 12, 1565-1581, https://doi.org/10.5194/cp-12-1565-2016, 2016.