We thank the reviewer for carefully reading our paper and our suggested changes. Please find our further changes below in italics.

Review of “The Sun's Role for Decadal Climate Predictability in the North Atlantic” revised by Drews et al. for publication in ACP

I’d like to thank the authors for considering my earlier comments and the effort they put into providing additional material. In my view, the paper has improved significantly. There are however still some issues which I’d like to see considered.

Page numbers refer to the version with highlighted changes.

I’m still a bit at odds with the partly inconsistent comparison to the study by Chiodo et al. (2019; C19 hereafter). In L193 the authors speak of similar correlations but different conclusions in comparison to C19. In L193 they say the C19 study “revealed” insignificant responses. Should this refer to the analysis of observations or simulations? If the latter the word “reveal” seems to be at conflict with the above statement. Furthermore, later the authors list many reasons for differences between the simulations of this study and Chiodo et al. As the aim is to “partly rebut the conclusions” of C19 it would be important to be very clear. Are the simulation results really different? Or is it just a different interpretation of similar results.

We refer to the correlations of the observed NAO index and the solar cycle. We now are more specific in the first sentence:
“Even though correlations for observations as in Fig. 5c are similar to what has been shown in a recent study (cf. Fig. 1b in Chiodo et al., 2019*), our conclusions are very different in the light of our model results.”*
We here mean that the simulation results are different as our experimental design is also quite different as elaborated (ensemble vs. single runs, more realistic CMIP6 forcing vs. idealized forcing, historical runs vs. constant GHG forcing). However, we explicitly do not rule out that similar results could be found in the C19 data if, e.g., monthly means instead of DJF means were analyzed. Since we do not have access to that data, we cannot verify nor reject this hypothesis.

I’m also confused by the new statement “seemingly discrepant results could be due to the analysis of DJF means in most studies, which likely are not sensitive enough to capture the signal reliably”. Does this refer to the papers cited in the bracket above "(Gray et al. ...)"? Why do they show signals if they used the not sensitive enough DJF means? If this refers to other studies, please cite.

We suggest to modify the paragraph:

However, as we show here, the solar signal is likely not present in all winter months. Therefore, in some studies there remains a signal in DJF means, while it could be averaged out in others, including ours (see below). This finding is also supported by the very recent study of Kuroda et al. (2022).

Furthermore, we adjust the text further down:

"~~While~~ We find statistically significant ensemble mean solar-cycle-induced surface signals in February during the strong epoch which are consistent with the top-down propagation from stratospheric signals. This limitation of the solar signal to February is also confirmed by Kuroda et al. (2022) using long observational datasets as well as a historical simulation with a different chemistry-climate model. A controversial ~~recent~~ study (Chiodo et al., 2019) ..."

L61 The authors contrast forced signals in the extratropical North Atlantic by saying that "up to 25% of decadal variance" is explained by the solar cycle and "this region shows low potential predictability due to other forcings. While this is not wrong it sounds like the solar influence is large compared to other forcings, which is not true. Averaging over this region by eye, I'd suggest that other forcings are still more important even in this region. It would be good to rephrase these sentences in order to avoid misinterpretation.

The reviewer is correct that this sentence has not been clear enough. We rephrased it to:

"The extratropical North Atlantic is a hotspot of solar cycle influence on climate predictability (Fig. 1a). In some parts, up to 25% of the decadal variance of winter surface air temperatures are explained by the solar cycle in our model. At the same time, similar parts of the North Atlantic show comparably low (<15%) potential predictability due to other (low-frequency) external forcings (Fig. 1b) and large internal variability (>65%) (Fig. 1c)."

We hope that this clarifies that this statement refers to only parts of the North Atlantic, and that it is similar regions that have a relatively large solar PPVF, low PPVF due to other external forcings, and large internal variability. We also added the numbers to the text for an easier impression while reading.

L80 “Consequently, solar variability and an adequate representation of its impact on climate is key to exploit the solar-induced potential predictability for decadal climate predictions.” This sentence sounds odd. Of course solar variability is key to exploit the predictability potential created by itself, and how could it be done in models if the representation of its impact was inadequate.

We understand the reviewer’s irritation regarding this sentence. We have to admit that various revisions in the process of writing the manuscript resulted in a too short description. We wanted to stress here that the exploitation of solar-induced predictability by means of a climate model requires a proper solar forcing as well as a climate model that is capable of simulating the whole top-down mechanism. This may seem trivial in the context of our study but we consider it worth mentioning that most existing climate model simulations do not match these requirements. Only recently, standard GCMs as used for the CMIP exercise started to employ spectral solar irradiance as forcing. A number of CMIP6 models still use only total solar irradiance which leads to an underestimation of the UV variability, which is the basic element inducing the top-down mechanism. When it comes to decadal climate prediction, there is - to our knowledge - no single prediction system including a model component for stratospheric chemistry. However, the ozone forcing recommended for usage in CMIP6 is compiled as three-dimensional fields evolving in time, including variability induced by the solar variability. Compared to earlier CMIP exercises, CMIP6 models hence may be able to model some response to solar variability even if not employing interactive chemistry.
We try to express that by rephrasing the respective sentence:

“Consequently, decadal climate prediction systems might benefit from including realistic solar forcing (e.g. SSI instead of TSI) and an adequate representation of its impact on climate (by usage of interactive chemistry or at least an ozone forcing matching the solar variability), exploiting the solar-induced potential predictability.”

L117 “This shows that the response to the solar cycle is highly non-linear and not necessarily proportional to the forcing.” Isn’t “non-linear: and “not proportional” the same? Why then the very different characterizations (“highly” vs. “not necessarily”?). I would agree with the latter but not the first statement. To show that the response is non-linear one would need to show that the response of the strong epoch scaled to the weak-epoch forcing is statistically significant different from the weak-epoch response.

We would like to refer to Figures S4 and S5, which show zonal mean temperature and wind during the weak and the strong epoch. In our opinion, it is clear from those plots that the response - compare weak and strong epoch - is not linearly scalable as suggested by the reviewer, i.e., the pattern is different, even reversed at times, and it is not the same pattern with reduced magnitude in the weak epoch as *compared to the strong epoch. However, we agree to align the nuances and suggest to delete "highly", while keeping the "not proportional" part as a clarification:*

"This shows that the response to the solar cycle is ~~highly~~ non-linear and not necessarily proportional to the forcing."

L158 "the correlations reach statistical significances of 90% in the model, and in observations with a lag of two years " Although the different lags are mentioned in the preceding sentence, this sentence on its own can easily be misunderstood. The different-lag issue should be picked up again. Besides, I don't think that one more line (as kindly produce in the response to my earlier review would make the figure too busy. So please include it. Furthermore, The correlations reach barely outside the 90% range. And, leaving aside autocorrelation issues, wouldn't one expect 10% of the values to be outside this range accidentally?

1. *The figure is now updated.*
2. *True, one would expect 10% of the values to be outside accidentally. Note that there are some values "significant" 1890 and 1910 (for the model), which seems to be rather accidental.*
3. *As a side note and described in the Methods, we took into account autocorrelation of the time series.*
4. *We suggest the following changes of the sentences:*

"Their running correlation for all overlapping 45-year windows is fluctuating in the earlier years but begins to rise in the 1920's both for the model (at 0 lag) and observations (first with a lag of 2 years, later 0 years) (Fig. 5c, cf. Fig. 7a in Kuroda et al., 2022, who also see a shift of the lag from +2 to 0 in observations around 1975 as shown here). In the second half of the 20th century, the correlations reach statistical significances of 90% ~~in~~ for the model NAO at 0 lag, ~~and~~while in observations significant correlations appear when the NAO lags the solar cycle by~~with a lag of~~ two years (Gray et al., 2013; Thiéblemont et al., 2015) and since the 1980's without lag"

L163 "organization and synchronization of internal variability" Is organization different to synchronization? Please explain. And I'd say the NAO index is synchronized, internal variability remains internal variability.

The reviewer is right that differences between "organization" and "synchronization" are marginal. Since we think that "synchronization" might imply a phase-locking to 0 lag, while "organization" does not, we prefer the term "organization" and delete "synchronization". We also agree that external forcing can only interfere with natural, and not internal, variability. Hence we now write "NAO":

“organization ~~and synchronization~~ of ~~internal variability~~ the NAO”

L193 The meaning of the “hence” is not clear to me.

We deleted “hence”.

L207 “The NAO in turn is organized and synchronized by the solar cycle. “

Not sure why “in turn”. Besides there is again the issue of organization and synchronization.

We rephrased this sentence to:

“We show that a stronger solar cycle signal induces a surface response that resembles the NAO, and the NAO-like index ~~in turn~~ is organized ~~and synchronized~~ by the solar cycle.”