Review of
Solar association with winter synoptic situations in the north Atlantic - European sector

This is an interesting paper that includes some careful analysis of a challenging topic. It examines the solar and QBO response in long-term observations of North Atlantic and European synoptic weather patterns. Because the synoptic patterns are relatively short-lived, of order days, there are many occurrences over a season and thus more data points than the monthly-averaged data that are usually employed in such studies. This helps improve the statistical significance of the results, which has always been a challenge for research on this topic. The authors conclude that there is a statistically detectable surface response to high levels of solar forcing compared to low solar forcing and this response is increased if only westerly phase QBO years are considered.

I am very supportive of this paper being published and my comments and questions below are aimed at helping to bring clarity to the paper and its results.

Main Comments

1. Is the analysis of the full period 1908-2014 really isolating the solar response, or is it aliasing the anthropogenic response?
   The aa-index is used as a measure of solar forcing and the authors stress that previous studies have failed to consider the impacts of long-term secular variations in solar forcing (because they employ sunspot number or F10.7 that consistently returns to zero during periods of solar minimum). I agree that there is an important difference between the various indices and it is very useful to compare them. And yet the long-term underlying variability in aa index is also a problem. As shown in figure 3, using terciles on a data record that has a long-term underlying variation means that the early part of the data record is classified as primarily low solar forcing and the later period is primarily high solar forcing. So, when you composite into high and low solar forcing terciles, how can you be sure that your results are due solely to differences in solar forcing and not mixed up with changes in anthropogenic forcing? The time evolution of greenhouse gas forcing is reasonably well characterised and is probably sufficiently different from the aa index evolution that it can be regressed out (and this is done in most previous studies that use regression techniques). Anthropogenic aerosol variations are more problematic though. Their time variations are more complex, with decadal- and centennial- scale peaks and troughs that are more challenging to distinguish from the long-term solar forcing evolution. There is also regional dependence of aerosol forcing, making it much more difficult to know how to characterise it. I believe this is the primary reason why previous studies have concentrated on trying to isolate the 11-year solar variations. I would like to see a discussion of these issues.

2. Analysis of sub-periods (section 3.1.2)
   Am I right in thinking that in the analysis of the sub-periods the terciles are defined using only data from those sub-periods? If so, the years contributing to the high and low solar index composites will be different from the analysis of the whole period and some years might even contribute to the opposite-sign composites e.g. the years around 1980 are included in the high solar composite for the whole period but might be included in the low solar composite for the shorter period 1958-2014. If I am mistaken and this is not the case then I don't understand how you can achieve a

reasonably well-populated low solar composite for 1958-2014 with so few years contributing to it (essentially only the last 10 years). More explanation and discussion of this would be helpful.
At this stage of reading the paper, and partly because of this issue, I also got confused about whether you are primarily examining the impacts of solar forcing on long (centennial-scale) timescales or on quasi- 11-yr timescales. Some discussion of this would be helpful, including the possibility that different feedback processes might operate on these different timescales (e.g. via ocean / ice feedbacks).

3. Like-for-like comparisons of aa-index versus SSN.
   There are several statements throughout the paper stating that the aa index gives higher statistical significance than the corresponding analysis using the SSN. This is not surprising. There have been a number of studies that demonstrate a peak response to the 11-yr solar variations in F10.7 / SSN indices at lags of ~3-4 years. This was initially reported by Gray et al. (2013) and subsequently followed up by a series of papers that explored this further, including proposed mechanisms for the lag that involve ocean feedbacks (e.g. Scaife et al. 2013, Andrews et al. 2015, Gray et al. 2016, Ma et al 2018). The Gray et al. (2013) is referenced in the introduction in another context but there is no discussion of the lagged response. A comparison of statistical significances using different indices is certainly a valid and useful aspect of the paper but I suggest that a comparison at only lag-zero is inadequate and potentially misleading.

4. Figure 5
   The discussion leading up to figure 5 is primarily based around exploring whether the small response to the QBO is because the QBO phase definitions pre-1958 are less reliable. So I was expecting to be shown distributions of just the QBO responses in the 2 different time intervals. Instead, figure 5 includes much more information and I found its complexity unhelpful. Table 2 shows that most of the responses are not statistically significant, so I would suggest that table 2 is sufficient to show most of the results and figure 5 could therefore be substantially simplified e.g. showing only the QBO or showing only the results that are meaningful (significant).

5. Have you considered repeating your analysis using daily mean sea level pressure data? I realise that others have already done something similar (e.g. Huth et al), but it would directly address (a) whether the use of monthly-averaged data raised by Salby is really an important issue or not (e.g. repeat the analysis using both daily data and monthly data and compare the results) and (b) it would make a comparative link to previous NAO-related studies that used SSN (thus excluding the long-term solar variations) and found a statistically significant solar response only in the southern node of the NAO over the Azores region e.g. do you find a similar response only over the Azores when compositing with the aa-index that includes the longer-term variations. (This is just a suggestion).

6. Some of the terminology and/or phrasing could be usefully adapted to improve the clarity of the paper, especially the use of ‘composites’ rather than ‘periods’ (details provided below).

Minor comments

Line 8 (Abstract): it would help the reader to add some text to clarify what timescales you are examining
Line 10: typo (flawns – flaws)
Line 22, 'numerous studies did claim detecting this cycle': English needs improving
Line 28: I don't think Kodera and Kuroda used the phrase 'top down'. Reference to a more general review of all the proposed top-down / bottom-up mechanisms could be useful here e.g. the Gray et al 2010 review in Rev Geophys or Haigh 2011: https://www.imperial.ac.uk/media/imperial-college/grantham-institute/public/publications/briefing-papers/Solar-Influences-on-Climate---Grantham-BP-5.pdf) or Lockwood https://link.springer.com/article/10.1007/s10712-012-9181-3.
Line 29 'The UV flux and energetic particles precipitation, both modulated by the sun, do affect the mesosphere and stratosphere, with changes in the stratospheric winds and the northern polar winter vortex.' A reference would be useful here.
Line 33 reference to Baldwin and Dunkerton – I think this may be the wrong paper reference, since it is not specifically about solar / QBO influences.
Line 54: there are many studies that cover more than 3-4 cycles using SST and mslp datasets back to 1850 and even using the Bronniman dataset that extends back to the 1600s (Gray et al. 2016).
Line 59 'which suggests that the length of records may be limiting the detection of very small signals': agreed, but an alternative suggestion is that models are unable to properly capture the correct mechanisms
Line 61 'found over recent decades is not found over earlier periods': it might be useful to reference the study of Ma et al 2018 Here; they use a sliding window and suggest that the response depends on the amplitude of the 11-yr solar cycle, which varies over time.
Line 62 'this suggests that either the correlation is found by chance or that the solar-climate relationship is not stationary': these are both possibilities but not the only ones; it is also possible that the response depends on the amplitude of the forcing, so that the mechanism may still be operating but the response is not detectable above the noise during periods when the forcing amplitude is low. Please also be aware that the Chiodo paper is based entirely on model data; I believe there is an invalid assumption in that paper, that because a model does not reproduce the signal then it cannot be real. I suggest being much clearer about whether these references are studies of actual data or models
Line 66: I believe most studies that only use short periods acknowledge that the lack of data is an issue - most surface studies use much longer datasets, only analysis of the upper atmosphere use short datasets because longer ones are not available. There is no need to justify your study by implying that all previous studies were flawed.
Line 69: you concentrate on examining the North Atlantic – European sector but give little background information on previous studies. It would be useful to include a few sentences here to describe the main results of previous studies of this region e.g. the Woollings et al., Gray et al., Huth papers are referenced at various points in the paper but there is no useful summary of their results, and there are also other relevant papers (listed below). It would also be useful to add a reference or two on studies that have examined the QBO impact at the surface e.g. a general review paper such as this very recent one https://doi.org/10.1038/s43017-022-00323-7 and/or the study of Gray et al. 2018 that examines the Atlantic QBO signal in more detail.
Line 84 'why it is so much variable': please re-phrase
Line 116: I'm not sure the word 'hence' is correct here.
Line 122 'as signed by a westerly flow': please re-phrase
Line 129 'all studies have used reanalysis outputs': this is not strictly correct, many have used much longer surface datasets including HadSLP

Line 133 'a solar-weather relationship could only arise from assimilating meteorological conditions': surely the dominance of the observations over the model is preferable, isn't it? Why is this considered a limitation?
Line 135: this reference to Baldwin and Dunkerton paper seems inappropriate to me - their paper is about stratospheric observations and not surface obs, which are much more frequent and reliable; it is also a very old paper and analyses have vastly improved since 1989.
Line 141 'it offers a rare opportunity to address solar-weather relationship at the daily time scale based only on observations rather than on reanalysis': yes, I agree with this statement but I was not convinced by the reasons you have given, which seem to rely on unfounded criticisms of previous studies
Line 191 'to numerically introduced': typo
Line 192 'Maliniemi et al. (2018) subtracted the long term trend to the SSN series. Since minima are almost constant, this detrending artificially transfers part of this long term trend from the maxima to the minima, making the SSN series more symmetric': you do not say whether you think this is good or bad, and if bad then why is it bad?
Line 194: I think it would be useful to show your 5-yr smoothing overlaid on this figure
Line 199 'The climatic mechanisms addressed with this index could be related to energetic particle precipitation.': this needs a reference
Line 215 'The associated largest changes in LWT have been found using a decadal smoothing of the aa index (averaged over 61 months with a Gaussian filter with a width of 10 months). Although synoptic activity has a typical time scale of few days, LWT changes can only be associated with long term changes in the aa index': surely LWT changes can occur on seasonal timescales?
Line 218 'This result is consistent with previous studies showing that solar activity at periods shorter than the 11-yr one cannot be detected in climatic observations': I got very confused here. What do you mean by 'long-term changes in the aa index' – do you mean quasi- 11-yr or centennial scale changes?
Line 220: 'provides us with six periods of time to compare:': most people would use the term 'composites' (or 'epochs') at this point, and I think it would help the reader if you use one of these terms. I also think it would be useful to state exactly which years go into which composites, especially given the possible confusion when you use terciles determined from different time intervals.
Line 235 'quantified by the difference in LWT proportions between two periods.':I think you mean composite differences here? Using 'two periods' is confusing, because it could be misinterpreted to mean two time intervals.
Line 273 'what could be expected': please re-phrase.
Line 281: I'm not sure this comparison of significance test with Anstey and Shepherd is required - although certainly this paper should be referenced somewhere in the introduction.
Line 286: The Camp and Tung study is worth discussing here, and also Gray et al. 2004 where they examined the influence of the solar cycle and QBO on the timing of SSWs. The Gray et al. study discusses a mechanism for this non-linearity in terms of the impact of solar and QBO on the timing of SSWs.
Line 336 'not necessary to associate the QBO phase to the aa index to detect a significant solar association with LWT,': I don't understand this sentence - none of the solar-only values in table 2 were statistically significant, it was only high solar + QBO-W combined that gave a significant response. And you yet you say it is not necessary to combine them?
Line 363: This is a rather surprising statement and could be rephrased; Labitzke et al never employed the aa-index, and of course they had far fewer years of data. I suggest that this could be re-phrased more positively; the work of Labitzke has been updated and now there is sufficient data for the signal to be significant when only the solar index is considered; the

impact is enhanced when both high solar and QBO-W are taken together, in agreement with Labitzke's study.
Line 369 'By construction, LWT cannot be influenced individually by different solar or QBO conditions, but only their occurrence and succession representing synoptic situations.': I don't understand this sentence
line 395: It would be helpful to also discuss this in terms of the NAO, and especially the strength of the Azores High, since most previous studies have highlighted a solar influence on the Azores High
Line 439: please see main comment on a fair comparison using lagged SSN index
Line 456: This seems contradictory, since northward means towards the north i.e. implying a meridional flow - do you mean a more zonal circulation positioned further north?
Line 470: please see previous comments on non-stationarity vs amplitude dependence e.g. the study of Ma et al. 2020.
Line 492 'relationship did claim significant association and conclude to solar impacts': please re-phrase.
Line 500: do you mean a large amplitude solar variability here? Can you clarify whether you mean that the background level needs to be larger, or whether you mean the amplitude of the 11-yr cycle needs to be larger.
Line 506: This is a vague sentence and I don't really know what point is being made here. 'Some studies have actually tried to include some secular solar variability' - please state which studies you are referring to; are you suggesting that this is somehow incorrect? I'm not sure.
Line 519: 'This relationship was found stronger using the aa index than with SSN, and this may point to different mechanisms implied in the relationship': I'm not convinced that this has been properly tested.

References

AA Scaife, S Ineson, JR Knight, L Gray, K Kodera, DM Smith, 2013
A mechanism for lagged North Atlantic climate response to solar variability
Geophysical Research Letters 40 (2), 434-439

MB Andrews, JR Knight, LJ Gray, 2015
A simulated lagged response of the North Atlantic Oscillation to the solar cycle over the period 1960–2009
Environmental Research Letters 10 (5), 054022

H Ma, H Chen, L Gray, L Zhou, X Li, R Wang, S Zhu, 2018
Changing response of the North Atlantic/European winter climate to the 11 year solar cycle
Environmental Research Letters 13 (3), 034007

LJ Gray, JA Anstey, Y Kawatani, H Lu, S Osprey, V Schenzinger, 2018
Surface impacts of the quasi biennial oscillation
Atmospheric Chemistry and Physics 18 (11), 8227-8247