"Solar association with winter synoptic situations in the north Atlantic – European sector"by G. Delaygue et al. subimmted to WCD
Comments of the authors to the reviews / wcd-2022-33-AC
16 Sept. 2022

 1. General comments
 2. Response to Community comment (Laken & Stordal 2016 paper)
 3. Response to Reviewer #1
 4. Response to Reviewer #2

----

1. General comments

We thank both reviewers for spending time to carefully read and comment our draft, which did not bring much anew but to explore what information could be extracted from a daily series.

We also want to apologize for some statements/ideas which have been overstated. We no where did write (nor thought) that "that all previous studies were flawed". We understand that our statements could be interpreted as "unfounded criticisms of previous studies", which was not our aim. Our aim was mainly to recall 1. that 'solar forcing' should be better defined, as different indices represent different aspects of the solar activity; and 2. that using seasonal averages may limit the statistical significance of correlation with climate series. These ideas were central in studies up to the 1990s and may have been forgotten, for possibly good reasons since the available records have lengthened, but should still be considered.

The point of Reviewer #2 on the role of other possible forcings (anthropogenic, volcanic, etc.) was very beneficial to us to try to clarify the problems posed by analysing a daily synoptic series. Considering (as proposed) a multiple regression shows that the different explanatory variables we could think of (solar, volcanic, anthropogenic) are able to explain a very small part of (the variance of) a daily synoptic record. The reason is probably simple: the atmospheric movements are partly chaotic, non-deterministic, at the weekly, synoptic timescale, so that most of its variance is unforced. The multiple regression technique was used to test the stability over time of correlations, in different ways (increasing the number of winters, moving a fixed time window, etc.). **These tests were very problematic, showing intervals with highly significant correlations alternating with intervals with insignificant correlations (Fig.1 below).** This instability is probably related to the very small part of the variance explained by the correlations, so that when this part increases a bit, the correlation is much higher, and its statistical significance is very high due to the very large number of data (12973 winter days over the full period 1908-2014).

This instability raises the question whether the correlations are actually due to the solar forcing, or could be produced with any signal.
Several lines of evidence:

       A. with the multiple regression we tested different explanatory variables in addition to a solar forcing (*aa* index or SSN): a long term increasing trend (anthropogenic forcing), a decadal volcanic signal (AOD), and noise (autoregressive process of 1st order): the first two are found mostly well correlated, whereas the last one is never found correlated.

       B. A correlation with LWT is found much better with an SSN series shifted back in time (leading LWT) by 1 year, and secondarily by 5 to 7 years (i.e., half the 11-yr period), with a slope reversing from negative to positive (cf. figure 2 below), which suggests that the correlation is (at least partly) related to the 11-yr solar cycle.

       C. A bootstrap resampling of the composite days, stratified by high or low solar conditions, produces some significant differences compared to the climatological means. (A bootstrap resampling is able to account for the heterogeneity of the series.)

Any advantage in using daily synoptic series (LWT) rather than seasonal averages (of Z500 etc.)?

D. Obviously the large number of data makes some results highly significant. There is some level of autocorrelation in the LWT series (with a lifetime of about 3.5 days), which reduces the number of degrees of freedom, but 1. even reducing the number of independent data by a factor of 3 would not change much the highest t-patameters and lowest P-values, and 2. a bootstrap resampling of composites gives the same standard deviation.

E. Daily weather types can be linked with synoptics at the basin scale (this is the reason of their definition): we could only illustrate how LWT successions can be related to typical synoptics, and how solar forcing may affect their probability, but a systematic relationship would need a dedicated work.

F. However the synoptic variability cannot be largely described by any simple forcing (the proportion of variance explained with a multiple regression is very low). By smoothing this variability, it is clear that regressions using seasonal averages are able to explain much more of the climatic variance, which mainly contains interannual and decadal (and longer) time scales. The drawback of using averages seems to be a decrease in the number of degrees of freedom and thus of the level of risk

The draft has been completely reworked, to: 1. present the multiple regression technique; 2. account for SSN ansd its strong association with LWT; 3. underline the contradiction between the very low uncertainty on regression slope and the lack of robustness of this slope.

Technically, the considerations about LWT successions (transition probability and transition changes) have been moved from Section 3.2 on 'LWT sequences' to Section 3.1 'Sensitivity of LWTs' where they justify the choice of both LWTs groups.

The feedbacks of the Reviewers, especially on the lack of robustness of the LWT association, would be very much appreciated in order to finalize a revised draft.

Figure 1. Slope and its P-value of the LWT regression by the *aa* index over an increasing interval, starting from 2014 on the left (regression over 1 winter = 120 days) and extending progressively over the whole 2014-1908 interval to the right (regression over 107 winters = 12973 days). (The *aa* index is one explanatory variable with 3 others: volcanic AOD, anthropogenic trend, and an AR1 noise.)

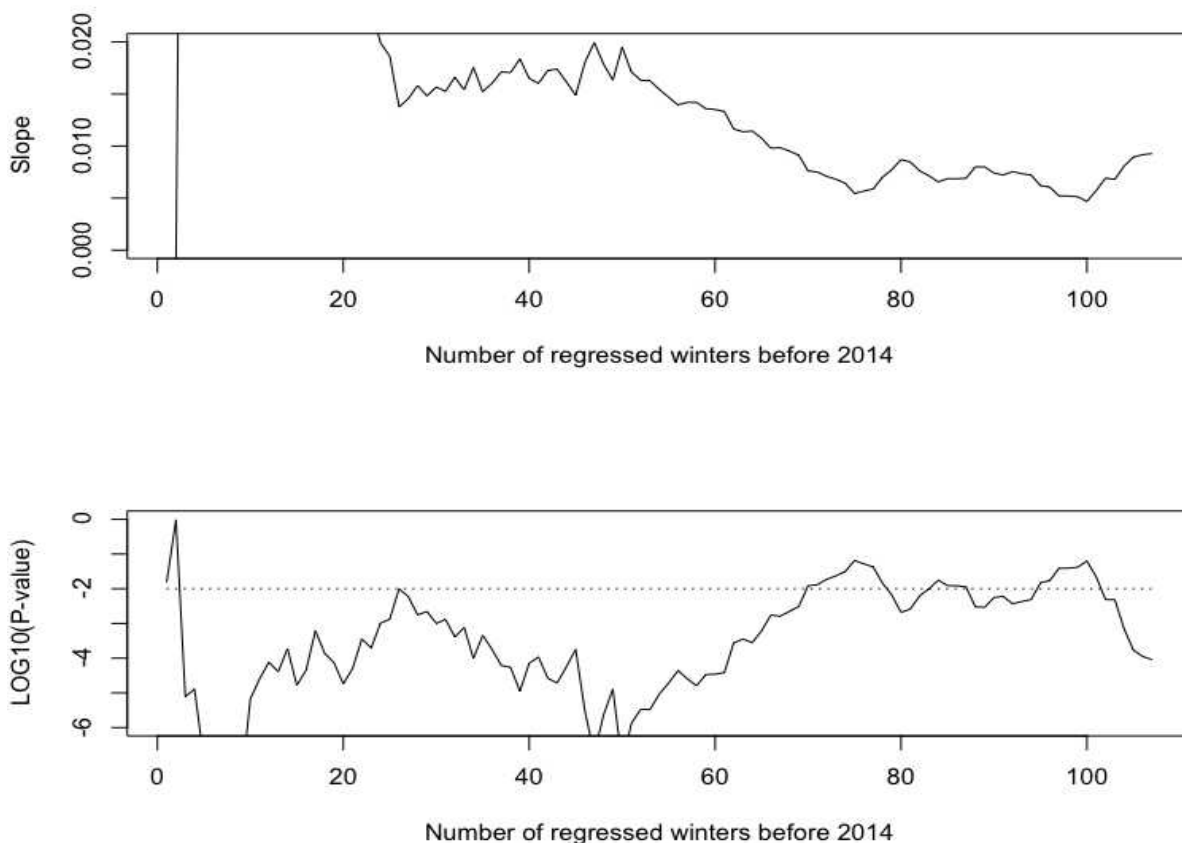
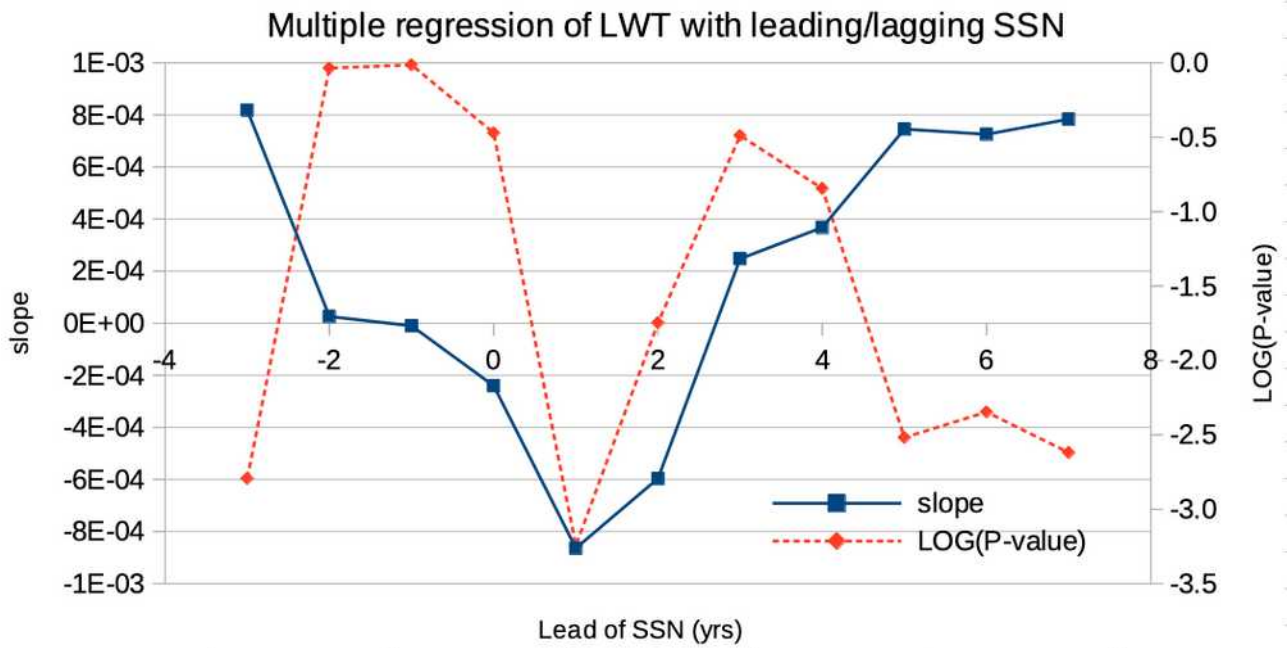



Figure 2. Slope and its P-value of SSN from multiple regressions of LWT with several forcings, over 1908-2014, for different lead or lag of SSN wrt LWT. The highest sensitivity and lowest P-value are found for a lead of 1 year (LWT lagging by 1 year). Secondary optima are found for leads of 5 to 7 years, or for a lag of 3 years.

2. Response to Community comment (Laken & Stordal 2016 paper)

Thanks for pointing this work, based on the Hess and Brezowsky Großwetterlagen series.
Huth et al. (2008) used this series as well, and i find this latter study more meaningful in terms of what can be inferred from this kind of data.

Review #1 of "**Solar association with winter synoptic situations in the north Atlantic** –
European sector"**by G. Delaygue et al.**

This is an interesting paper, which is worth publishing. Its main idea is that some types of atmospheric circulation occur with an enhanced / lowered frequency under specific phases of solar activity, quantified by aa index, and quasi-biennial oscillation. An interesting idea is an attempt of the authors to reflect the secular component of solar activity; it is typically neglected in analyses of solar-climate links, which in their majority use characteristics of solar activity that describe 11-yr cycle well (sunspot numbers, F10.7), but not its long-term trends. While the paper does not bring breakthrough ideas and findings, its main virtue is a carefulness of statistical analysis.

Thanks for these comments. One goal of our study is to make it clear that sunspots, F10.7, heliomagnetic changes etc. are different expressions of the solar activity (which originates in its internal magnetic field, see, e.g., Lockwood, Stamper and Wild 1999; de Jager 2005). Instead, many studies consider the 11-yr cycles (with a sunspots series) as the only 'solar activity'. These expressions of the solar activity have different time scales and different potential climatic forcings (L.158 ff). The variability of the aa index is dominated by a secular component, and its potential climatic forcing is the modulation of high energy particles flux; both aspects are interesting to remember. It is actually worth underlining that some studies based on sunspot or F10.7 series did modify them (e.g., with normalisation: L.191-194) in order to increase their correlation with climate, with the consequence to decrease the 11-yr component and increase the secular component. Hence one message we have is that studies should make it clear which part of the solar activity is accounted for, rather than which solar series is used.

Nevertheless, there are some aspects of the analysis and statistical evaluation, which are not described with enough precision and ask for clarification, or are to some extent questionable:

- The aa data are smoothed; smoothing always lowers the number of degrees of freedom. The smoothing procedure should be better described and the effect of smoothing on the reliability of results should be discussed. Were the smoothed data interpolated back to daily resolution, as l. 211 suggests?

The smoothing technique was described at L.216: "averaged over 61 months with a Gaussian filter with a width of 10 months". Not sure what could be added to this description? And yes, since our climatic data have a daily resolution, the smoothed aa series was daily interpolated, as stated at L.211.

Smoothing lowers the number of degrees of freedom in the forcing, but not in the LWT series which is used for composites and regression. This smoothing of the aa index has been abandonned in the revised draft, since the regression slope was found slightly more robust with a monthly *aa* index and SSN.

- The synoptic types are merged into two groups *a posteriori*, only after their association with solar activity is evaluated: types with enhanced frequency under high solar activity (A, W, NW) and under low solar activity (C, N, E, S) are put together and the association with solar activity is examined for these groups. This may degrade assessments of statistical significance. Although some justification for both groups is provided, I do not find it convincing enough, particularly for the inclusion of the A and C types.

This association comes from both the definition and simplification of LWTs. LWT is the dominant flow direction or high/low pressure system over a small area. We further simplified this description to the main directions of flow. (Similarly, the GrossWetterTypen are simplifications of the GrossWetterLagen). Hence any day, with one specific LWT, cannot constrain the synoptic situation over the north Atlantic-European region; rather it is the succession of LWTs which constrains the synoptic situation and development. Statistical analysis showed how LWTs are strongly associated,

into 2 groups: A, N & W on one side, and C, N, E & S one the other. In Section 3.2 we tried to show that these associations do constitute typical synoptic trajectories.

Linking these LWT associations to synoptic situations would probably be clearer by using the whole set (26) of LWTs, producing more realistic synoptic successions. For instance, instead of associating A to N & W, we would probably find AN associated to N & W. However, we would face a much larger diversity of LWTs successions (potentially $26^5 \approx 12E6$ over 5 days) which would be practically impossible to analyse.

Further general comments:

While aa index is certainly related to solar activity, it is primarily descriptor of geomagnetic activity. Hence the focus of the paper on solar activity, advertised in the title and throughout text, may be found confusing or misleading by a reader.

The aa index is a geomagnetic index by construction, but over the past 50 years studies have shown that it is indeed a quantitative proxy of the heliomagnetic field variability (see, e.g., Clilverd et al. 2005).

On the contrary, the long and impressive series of sunspots number is not a quantitative proxy of any 11-yr cycle feature, especially not of total irradiance. This is because sunspots are very localised magnetic anomalies, which do affect the photospheric emission but do not dominate its 11-yr variability.

The analysis of 5-day sequences of synoptic types is an interesting approach; it is inconclusive and does not provide any relevant result, however. Therefore, it may be entirely omitted or at least substantially reduced.

Why would this approach be 'inconclusive'? On the contrary we think it is a very valuable way to interpret statistics of individual LWTs as synoptic changes. The result that one group of LWTs have been more/less frequent under different solar/QBO conditions is of limited interest; however we showed that the 2 groups of LWTs correspond to contrasted synoptics and were able to interpret changes in LWT frequencies as changes in synoptics.

What is correct is that we are not able to analyse all 5-d sequences in terms of synoptics because they are far to many (~3000). This is a limitation related to the static (local) definition of LWTs. (Tracking the synoptics would require to analyse, e.g., the Z500 field covering the whole European-north Atlantic area; see for the particular example of blockings, e.g., Barriopedro, Garcìa-Herrera & Huth 2008). The static/local definition of LWTs implies an approach of LWTs succession in time (rather than in space, as with analyzing a large field).

What we could do is to diagnose changes in the lifetime of individual LWTs, changes which make LWTs more or less frequent under the different solar/QBO conditions.

Specific and minor comments:

Introduction: You often talk about 'correlation' of solar and climate variables. I am not sure if you really mean only correlation (then, why you dismiss those many studies using other tools than correlations) or an association in general (then term 'association' or similar may then be more appropriate).

Correct, our intention was to use 'association' instead of 'correlation'; this has been corrected.

Ll. 54-55: The situation described by Salby and Shea in 1991 (that analyzed records typically covered 3 to 4 solar cycles) was true thirty years ago, not today.

Unfortunately, some recent studies still use short periods or few points to do statistics:

Roy 2018 (published in *Nature Scientific Report*) used the NCEP/NCAR reanalysis over the period 1979-2016, covering 3 and half solar cycles (4 maxima and 3 minima; see fig.1a);

Laken and Stordal 2016 used SSN over the period 1881-2000 but considered only the solar maxima to composite climatic data, hence doing statistics with 11 values;

Anstey & Shepherd 2014 used the ECMWF reanalysis over the period 1957-2011, covering 5 solar cycles;

Roy & Haigh 2011 used the NCEP/NCAR reanalysis over the period 1953-2004 covering 5 solar cycles;

Frame & Gray 2010 used the ERA-40 reanalysis over 1979-2001, extended by ECMWF operational analysis to 2008, hence covering 3 solar cycles;

The historical excursion on ll. 84-89 (and partly also in following paragraphs) is a bit redundant and remote to the focus of the paper.
ok, removed

L. 96: The origins of Hess-Brezowsky classification go further back in history, at least to 1952.
Reference changed to: HESS, P.; BREZOWSKY, H.: Katalog der Großwetterlagen Europas. Berichten des Deutschen Wetterdienstes in der US-Zone, Bad Kissingen, 33, 1952

L. 109: 2016, not 2014.
updated, thanks (i add an earlier version of it)

L. 125: 'Mode of circulation' may associate with teleconnections (modes of low-frequency variability), which I believe you do not have in mind here.
Yes modes as teleconnections, defined from PCA of, say, Z500 or SLP fields. Instead weather types have been defined to define synoptic situations.

L. 129: 'all studies' is too strong an expression: I am aware of studies that do not use reanalysis data.
Correct, 'all' only referred to the studies referenced before, so this was incorrect [changed].

Text in ll. 132-134 lacks precision: Which reanalyses are 'classically used'? It is the models used to produce reanalyses, not the reanalyses themselves, that 'do not account for solar cycle forcing'. Furthermore, this claim is at variance with text in parentheses about 20CR and ERA5 (which do account for solar cycle).
'reanalysis' is in fact a loose term since it describes the system of observations analysis as well as the outputs thermselves.
Yes, 'classically' is an incorrect term, rather 'mostly used' for NCEP (i guess 20CR or ERA5 have not been so often used as NCEP?).
The part of the study concerning the use of the original, 'subjective' LWT series, has been removed (and the discussion of reanalysis uncertainty).

A logical jump is on l. 134 from representation of solar signal to (dis)agreement among reanalyses: a start of a new paragraph should be inserted in front of 'this limitation'.
[paragraph removed]

L. 136: Which 'series' do you have in mind?
[paragraph removed] Hanson et al 2004 did compare the NCEP and ECMWF reanalysis.

L. 150: The basics of the 'simplification scheme' can be explained here.

L. 189: 'regress' may not be enough general a word here

Changed to 'stratify' since this is the case described here.

L. 192: subtracted from
thanks!

L. 199: The claim about association of aa index to EPP should be supported by a reference.
Sinnhuber & Funke 2020 ?

L. 211: How was the interpolation done?
linearly between monthly values

Term 'change' in section titles (3.1, 3.1.1 etc.) and text may be confusing as it associates with a process in time. 'Difference', 'association' etc. may be more appropriate.
thanks; section titles have been modified (Section 3.4), and the text throughout

Most of text on ll. 236-247 is a description of methods and may be more suitably placed in Sec. 2.
simplified and placed at different points

The rationale for bootstrap resampling and ways how it was conducted should be described more clearly.
It has been reworked; hope it is clearer; the technique is simple, though.

L. 255: is not clear, 'different' from what (twice on the line)
Difference if each LWT group proportions between the intervals; I've added A, W & NW LWT to make it clear.

Claims on ll. 273 to 275 seem to contradict each other ('combining QBO and solar conditions does NOT lead to a change in LWT occurrence' versus 'changes are much stronger by combining W-QBO to high solar conditions').
It was written "not systematicaly lead", i.e., in all configurations.

Why are three or four different climatologies shown in Figures 4 and 5?
[figures removed] Climatologies were recalculated from the whole interval with the same number of data as each composite, to get a more realistic estimate of its uncertainty (which is stronger over a shorter interval). Note the same mean but larger spread of these climatologies calculated over shorter composite.

The legend obstructs parts of graphs in Figs. 4 and 5, which are relevant – please move it away.
[figures removed]

I believe the black distributions (climatology) should be the same in the left and right graphs: why are they not?
[figures removed] They have the same statistical distribution, but are not exactly similar since issued from different resampling.

L. 283: Please be consistent in using abbreviations.
modified

L. 287: This also seems to contradict the claim on l. 273.
We use monthly *aa* index and SSN in the revised draft, and the association of QBO with a solar index is associated with even smaller differences than with smoothed indices.

L. 315: Fig. 5d
thanks [figure removed]

What is the end of the first period on Fig. 5: 1957 (as shown in the heading) or 1961 (in figure title)?
[figure removed]

The green distributions (SSW) in Fig. 5 right are not discussed; only a brief mention without reference to this figure is provided in the discussion.
discussion removed so far, may be reintroduced if still important

I believe that 'LWTs' should be written when used in plural (e.g. l. 321).
yes, I've tried to adopt LWTs specifically when it comes with 'different LWT' rather than LWT as a series.

Ll. 334 and below: A possible, although not likely, alternative explanation may be that atmosphere responds to phases of the solar cycle rather than to the magnitude of the forcing.
Not sure which part of the text this refers to. By solar phases, does it mean increasing and decreasing phases? That is possible; however since these phases are quite periodic we could expect a strong climatic signal around 11-yr, almost not observed, is that correct?

Ll. 363-365: You should consider here that Labitzke et al. did not work with aa index, which is a likely reason for the difference.
That is true [modified]

Some clarification is needed of term 'synoptic situation' (l. 369 and below) and how it differs from 'synoptic pattern', 'synoptic type', etc. For me, for example, terms 'situation' and 'pattern' are almost equivalent.
Yes, 'situation' as 'pattern'. Does it make sense or pattern is better/clearer?

L. 373: 'These sequences are used to explore the association of LWT found in Sec. 3.1' – but individual daily LWTs, not sequences, are explored in Sec. 3.1.
[removed]

Ll. 373-4: 'also to which synoptic situations correspond these sequences' – something is wrong here.
[removed]

L. 376: 'LWTs are not randomly associated': (i) 'composed' may be more appropriate here than 'associated'; (ii) but the randomness of composition of sequences is substantially limited by their overlap.
which overlap? LWTs are simplification of a pattern, so by definition they have no overlap (but the patterns have, of course).

L. 381: 'strongly' – is the link within the two groups of types really strong? Its strength should be quantified in terms of statistical significance: how (un)likely is such a link?
This is difficult to address other than by the frequency of transitions (this discussion has been moved to Section 3.1). Another index is the transition frequency normalised to the frequency of each LWT, to account for the strong differences between LWTs. However, for the rare LWTs (E, S) this normalisation probably leads to overestimate the LWTs associations.
Other groups of LWTs could probably have been considered. The two groups considered in this work (A, W, NW and C, N, E, S) come from empirical considerations of the transition frequencies and their individual sensitivity to solar index.

Ll. 382-3: Causality is not obvious for the last part of the sentence.
Ok, removed

L. 389: As one example of situations related to the C+N+E+S group, sequence ASCCC is given; but A is from the other group.

A and C LWTs by far dominate, S is very rare but usually associated to synoptic change with the progression of a low (C) displacing a high (A); i think this sequence ASCCC is interesting as typical of weather change.

L. 390: ASCCC (too many C's)

thanks

L. 394: Reference to Tabs. 1 and 2 is confusing here as the tables show frequencies of individual LWTs, not sequences, which are discussed here.

[removed]

L. 413: 'significant' in a statistical sense?

yes in a statistical sense, i.e. from the associated uncertainty

'Shortening of Cyclonic persistence' (l. 427) is a likely explanation, but an alternative is also plausible: there may be fewer C sequences without changing their duration.

The question is why fewer C LWTs: since we consider successions/chains of LWTs, fewer means either that C was less often created or more often replaced, hence this can be quantified with transition frequencies.

L. 449: add 'in its positive phase' after 'NAO pattern'.

Done

Discussion on blockings: you may also refer to Barriopedro et al. (2008, J. Geophys. Res. 113, D14118) who investigate the association of solar activity with blocking anticyclones.

see above

L. 462 and below: you may add to the discussion the paper by Huth et al. (2008, Ann. Geophys. 26, 1999-2004) where the association of Hess-Brezowsky types with solar activity is presented.

Done (i knew only the EGU presentation but not the corresponding published paper.)

L. 476: 'primarily' may be too strong here: There are other effects that modulate the strength of the polar vortex than SSWs. Actually, SSWs temporarily destruct the polar vortex rather than modulate it.

'Primarily' as the most important source of variability; but it is correct that SSW does not 'modulate' the vortex (since modulation rather implies continuous modifications). Changed to "very strong disturbance" of the vortex.

L. 479: 'they should hold with SSW' is unclear.

[removed]

Please check correct spelling of names with diacritic – both in text and the reference list. Some references are incomplete: volume number and/or pagination are missing.

Sorry for that; import of reference is not always perfect

References

Anstey and Shepherd 2014. High-latitude influence of the quasi-biennial oscillation, *Quart J Royal Meteorol Soc*, 140, 1–21, https://doi.org/10.1002/qj.2132

Barriopedro, Garcìa-Herrera & Huth 2008. Solar Modulation of Northern Hemisphere Winter Blocking. J Geophys Res 113, D14118, 10.1029/2008JD009789

Clilverd et al. 2005. Reconstructing the long-term aa index. *J Geophys Res* 110, A07205, 10.1029/2004JA010762

de Jager, 2005. Solar Forcing of Climate. 1: Solar Variability. *Space Science Reviews*, 120: 197-241, 10.1007/s11214-005-7046-5

Frame and Gray 2010. The 11-Yr Solar Cycle in ERA-40 Data: An Update to 2008. *Journal of Climate* 23, 8, 2213-2222, 10.1175/2009JCLI3150.1

Hess, P. and Brezowsky, H.: Katalog der Großwetterlagen Europas. Berichten des Deutschen Wetterdienstes in der US-Zone, Bad Kissingen, 33, 1952

Laken and Stordal 2016. Are there statistical links between the direction of European weather systems and ENSO, the solar cycle or stratospheric aerosols?, *Royal Society Open Science*, 3, 150320, https://doi.org/10.1098/rsos.150320

Lockwood, Stamper and Wild 1999.: A doubling of the Sun's coronal magnetic field during the past 100 years, *Nature*, 399, 437–439, https://doi.org/10.1038/20867

Roy 2018 Solar cyclic variability can modulate winter Arctic climate, *Sci Rep*, 8, 4864, https://doi.org/10.1038/s41598-018-22854-0

Sinnhuber, M. and Funke, B.: Chapter 9 - Energetic electron precipitation into the atmosphere, in: The Dynamic Loss of Earth's Radiation Belts, edited by: Jaynes, A. N. and Usanova, M. E., Elsevier, 279–321, https://doi.org/10.1016/B978-0-12-813371-2.00009-3, 2020.

von Storch & Zwiers 2002. Statistical analysis in climate research. Cambridge University Press

Review #2 of "**Solar association with winter synoptic situations in the north Atlantic** – **European sector**"**by G. Delaygue et al.**

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.

Many thanks for this point: i was too optimistic with the very high level of significance of composites differences.
To address this question of 'confounding variables', a multiple regression was applied to LWT, with different explanatory variables. The logistic regression applies to probability of daily LWT, and alleviates the need of using seasonal occurrence of LWT. This regression approach shows that solar or anthropogenic forcings could statistically explain only a very small part of the synoptic LWT variance, although significance of regression slopes is very high (see the Comment introduction for a more detailed description). In short: yes it is possible to separately account for the *aa* index and an anthropogenic trend (and they are both significant), but it is not yet clear to me what can be concluded from this regression.

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).

Sorry this was not clear. The same terciles values were used for all intervals, for consistency. Tables 1&2 of the first draft give the number of days used for the statistics: the minimum is 483 days (equivalent to about 4 winters).
Note that Figure 3 comparing the *aa* index to SSN is a bit misleading since showing the series over 1868-2018, a longer interval than the one analysed (1908-2014), so that the terciles values used for the figure are slightly lower than the ones used for our satistics, since the *aa* values over 1868-1908 were lower than after.

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.

True! With the multiple regression, we tested lead/lag of SSN, and the statiscal significance of slope is much higher for a lead of 1 yr (wrt LWT), and then between 5 to 7 years (see Figure 2 of our Comment introduction).

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).

Ok. The purpose of Figs 4&5 was to illustrate what means a *t*-parameter of 3 to 4 in terms of difference significance. I've skipped these figures.

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).

An interesting idea. Test whether the SSN pattern in a 2D, lon-lat, daily SLP series, is the same as the monthly/seasonal pattern? For the very same reason as for LWT, I'm wondering what could be a SSN pattern with daily SLP, dominated by synoptic variability. Typical modes like NAO are calculated with monthly/seasonal averages in order to filter out synoptics variability.

Also, this would be a different work: SLP is a continuous variable, and its daily variability is probably less directly linked to synoptics situation than LWT successions. The idea of using LWTs is that they represent simplifications of daily weather patterns, which successions define synoptics situation.
I cannot find work by Huth et al. with daily SLP. The group at Institut de physique du globe de Paris (IPGP) did examine daily temperature and SLP series (e.g., Le Mouël et al 2009 10.1016/j.jastp.2009.05.006). (However, since their aim was openly to show that the global warming was due to solar activity, their works have been somehow discarded. See for instance Legras et al. 2010 10.5194/cp-6-745-2010)

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).
many thanks (changed)

Minor comments
Line 8 (Abstract): it would help the reader to add some text to clarify what timescales you are examining
ok

Line 10: typo (flawns – flaws)
thanks

Line 22, 'numerous studies did claim detecting this cycle': English needs improving
ok

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.
Correct, still, Kodera & Kuroda 2002 paper is about "downward control principle", i.e., explaining "tropospheric effects through changes in vertical propagation of planetary waves".
Added Gray et al 2010.
Mike Lockwood's 2012 Review cited for his following sentence: "The academic reputation of the field of Sun-climate relations is poor because many studies do not address all, or even some of, the limitations listed above."

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.
Sinnhuber & Funke 2020 ?

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.
ok

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).

Correct. But there are still recent, and highly cited studies using no more than 3-4 cycles (Roy 2018, for instance, based on 3.5 cycles 10.1038/s41598-018-22854-0; also note that Laken & Stordal 2016 used only 11 values corresponding to composites based on solar maxima 10.1098/rsos.150320).

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
Correct

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.
Thanks for this reference, very interesting in fact.

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
Correct, this is what Ma et al 2018 suggested. But if "the response depends on the amplitude of the forcing", isn't it that "solar-climate relationship is not stationary"?

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.
We nowhere wrote, nor thought, that "previous studies were flawed"; a probably better view is that claims of previous studies are "on thin ice" (van Oldenborgh et al 2013 10.1088/1748-9326/8/2/024014) meaning that statistics are not very strong nor very clear.
I'm not sure that "most studies acknowledge that the lack of data is an issue" (otherwise "The academic reputation of the field of Sun-climate relations is poor because many studies do not address all, or even some of, the limitations listed above."

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.
That is entirely correct, however

Line 84 'why it is so much variable': please re-phrase
[paragraph removed]

Line 116: I'm not sure the word 'hence' is correct here.
[removed]

Line 122 'as signed by a westerly flow': please re-phrase
done

Line 129 'all studies have used reanalysis outputs': this is not strictly correct, many have used much longer surface datasets including HadSLP
correct!; SLP and weather types like HGW
[This paragraph as been removed]

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?
[This discussion as been removed]
But what happens if the solar forcing of the model and the assimilated observations were not consistent?

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.
ok

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
Reanalyse is a fantastic and powerful technique, the point is that reanalyses are usually considered as observations without any uncertainty. (Practically, uncertainties of reanalyses are very difficult to appreciate, but they do exist.)
The discussion on the original, subjective, LWT series has been removed because the results with the objective series are in themselves difficult to analyse.

Line 191 'to numerically introduced': typo
thanks

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?
"Good or bad" is not the point, such detrending modifies the forcing to a point that it is a completely different series, not SSN any more. This should be underlined, and if possible, a comparison would be given of the impacts without and with such modification (otherwise it is very difficult to compare the results of different studies).

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

Sinnhuber & Funke, 2020?

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?

Sorry the meaning was not clear here: solar forcings can be used at different time resolutions, the aa index is defined at the 3h resolution, SSN exists at the daily resolution. However it is probable that they have no climatic impact at these time scales, because the climatic system is not sensitive at these (short) time scales. One question is the time resolution at which some impacts may be detected.

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?

Sorry, 'long term' wrt the original definition of the aa index at the 3 hours resolution. (modified)

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.

many thanks ('composites' used throughout)

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.

yes, composite differences (changed)

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.

It is done since this study is conducted with daily data, so that statistics are done on hundreds to thousand of days. A very rare case to compare with.

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.

ok, thanks

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?

It depends on the level of risk accepted to consider a difference significant. In Table 2, the composite based on low solar conditions over 1908-57 has a difference to the mean with t-

parameter of 3.6, which can be considered as significant (with a Gaussian distribution, the probability above 3.6 is about 2E-4).

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.
Yes that is totally right; thanks for re-framing the context.

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
Sorry: a LWT is a diagnostic of the daily flow; the diagnostic cannot be influenced by a forcing, but its probability (occurrence) or the succession of LWTs (particular synoptic development). [suppressed]

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
The SLP pattern associated with the westerly LWT (Figure 1) is almost exactly a NAO+ pattern. Hence the strentgh of teh Azores High could be directly related to the proportion of W LWT.

Line 439: please see main comment on a fair comparison using lagged SSN index
yes cf. main comment

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?
a crest of high pressure developping towards the north

Line 470: please see previous comments on non-stationarity vs amplitude dependence e.g. the study of Ma et al. 2020.
Ma et al 2018 ? cf. main comment

Line 492 'relationship did claim significant association and conclude to solar impacts': please re-phrase.
ok

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.
Good point! The multiple regression applied over varying intervals shows that the association of LWT with SSN is strongest in the first half of the 20th century, and the association with the *aa* index strongest in the last part of the 20th century (based on the P-value of the regression slopes). But comparing both intervals i do not think it possible to decipher whether it is about background level or amplitude of the 11yr cycle.

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.

It referred to Maliniemi et al. (2018) study, in which the SSN series was modified by removing its long term trend, thereby introducing a a secular varibility.

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.
yes, SSN with a 1yr lag has also some association with LWT, so in terms of mechanism it is not yet clear.

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