Ann. Geophys. Discuss.,
https://doi.org/10.5194/angeo-2019-156-AC1, 2020

Interactive comment on "A note on the statistical evidence for an influence of geomagnetic activity on JRA-55 northern hemisphere seasonal-mean stratospheric temperatures" *by* Nazario Tartaglione et al.

Nazario Tartaglione et al.

nazario.tartaglione@norceresearch.no

Received and published: 12 February 2020

Major analysis/presentation comments:

(C1) The authors focused on point-wise analysis of temperature data and ultimately found no significant local signal related to geomagnetic activity. I wonder, however, if such signal could be more clearly seen in data averaged over larger areas, i.e. obtained for individual sectors, latitudinal bands, or over the entire extratropical area. I base this (possibly unfounded) suspicion on the presence of uniformly positive anomalies across

large segments of the analysis area, notable especially for the JJA and SON seasons at the 5 hPa level (Fig. 2). By averaging the temperature series from multiple grid points,signal-to-noise ratio can perhaps be improved; conditional averages considered by the t-test may then more clearly reflect the geomagnetism influence.

Answer: As the referee 2 made a similar comment, we will add a figure with the zonal average (FIGURE 1). Averaging zonally and applying both the corrections we do not have any significant area.

(C2) The autocorrelations seem to be substantial in some of the time series. The authors address their effect through a correction reducing the number of degrees of freedom considered in the t-test. Is there, however, any identifiable source of these autocorrelations (such as a long-term trend, or imprints of solar activity variations)? If so, removal of the respective components from the time series may potentially result in higher (and statistically more significant) contrast between temperatures pertaining to low/high geomagnetic activity periods.

Answer: In general, autocorrelation is mainly due to persistence of temperature patterns year by year. For instance this is the case of the large value of temperature autocorrelation found during the summer season. However local higher autocorrelation values during other seasons can also be due also to the low frequency variance caused by large scale teleconnections (see for example Madden, 1976). The Zwiers and von Storch method is considered a standard procedure to deal with temporal autocorrelation.

(C3) To quantify and visualize presence of autocorrelations in the temperature data, statistic of the Durbin-Watson (DW) test is shown in Fig. 1. Maybe presenting the lag-1 autocorrelations instead of (or in addition to) the DW statistic would better illustrate the autocorrelation structures, as they are directly involved in calculation of the corrections applied in the paper (eq. (1)), and arguably more intuitively interpretable than the values of the DW statistic itself.

Answer: We use the Durbin-Watson because it rules out the possibility that sampling error can cause the autocorrelation. If we instead would have used lag-1 autocorrelation we would need to check whether the lag-1 autocorrelaton is statistically significant, a task that DW test does automatically. However, the plot of the lag-1 autocorrelation looks like to that obtained with DW (FIGURE 2).

(C4) A requirement of Gaussianity is mentioned with regard to the t-test (l. 88), but, unlike other test assumptions, it is not tackled any further. I assume that this assumption is reasonably well satisfied, considering consistence of the data with AR(1) model (as discussed in the paragraph at l. 104+), but perhaps this could be mentioned explicitly?

Answer: We checked the normality of the distributions over the domain and yes, in general, this condition is satisfied in most of the area. However it is widely but incorrectly believed that the t-test is valid only for normally distributed outcomes. Efron (1969) for example has shown that the t-test is still valid under weaker assumptions. We will specify this point in the revised text. point.

(C5) For better comparability with topically close studies (especially Seppälä et al.(2009), by which much of the methodology in the current manuscript seems to be inspired), maybe results for lower atmospheric levels could also be shown/mentioned.

Answer: We agree with the reviewer and we will show the 2m temperature (FIGURE 3). After the two correction steps, there is still a small area with statistically significant difference over the Scandinavia. However, this only tells us that by this feature (warm Barents Sea) the selected years are unusual. The lack of a recognizable link with stratospheric anomalies makes it difficult to establish that it is related to the geomagnetic activity.

(C6) Fig. 2: The positions of grid points with statistically significant negative temperature differences (and their corresponding purple outline) seem suspicious: instead of being located within the areas pertaining to negative differences, they appear near the line separating the + and – regions

Answer: Thank you for this observation. There was indeed a small bug in the plot subroutine that produced such features. The new figures do not have this problem as we use gray shading for significant, at the 5% level, temperature differences (FIGURE 3).

Minor/technical remarks:

l. 19: “and is thereby” to “and are thereby

”Table 1: “2001” misspelled as “20001”

l. 74-75: Did Seppälä et al. (2009) really use daily-step data in their analysis?

Answer: Yes; we guessed it when we looked at their figures, and it was confirmed by a personal communication.communication.

l. 85: maybe reference to Benjamini and Hochberg (1995) would be preferable here, as they are the original authors of the FDR method (as discussed later in the manuscript)

We agree with the reviewer.

l. 144: Shouldn’t there be J rather than i in the numerator of the fraction?

J is the index representing the max value of the sorted p(i) values . We will rewrite the statement to make it clearer.

l. 146: Welch’s variant of the t-test is mentioned (i.e., the form assuming un equavariances of the samples compared), yet t-test employing pooled variance is presented earlier in the text (eq. (2))

Answer: The Welch’s test is always applied, even when analysing the original temperatures with n there is no correction. What is specified in equation 2 is the correction of the t-test with the Zwiers and von Storch method. In this case the pooled variance is computed with the equivalent values of n and m.

Fig. 3: The green outlines seem to be only partially drawn. 163: “point” to “points”

Answer: We solve this issue using a gray area (see Figures 3).

l. 190: extra comma

Answer: We thank the reviewer for these corrections. Most of the typos occurred when the text was converted from word to latex and they will all be corrected in the revised manuscript..

References

Bradley Efron (1969) Student's t-Test Under Symmetry Conditions. Journal of the American Statistical Association.Vol. 64, 1278-1302

Madden, R.A., 1977: Estimates of the Autocorrelations and Spectra of Seasonal Mean Temperatures over North America. Mon. Wea. Rev., 105, 9–18, https://doi.org/10.1175/1520-0493(1977)105<0009:EOTAAS>2.0.CO;2

Interactive comment on Ann. Geophys. Discuss., https://doi.org/10.5194/angeo-2019-156, 2019.

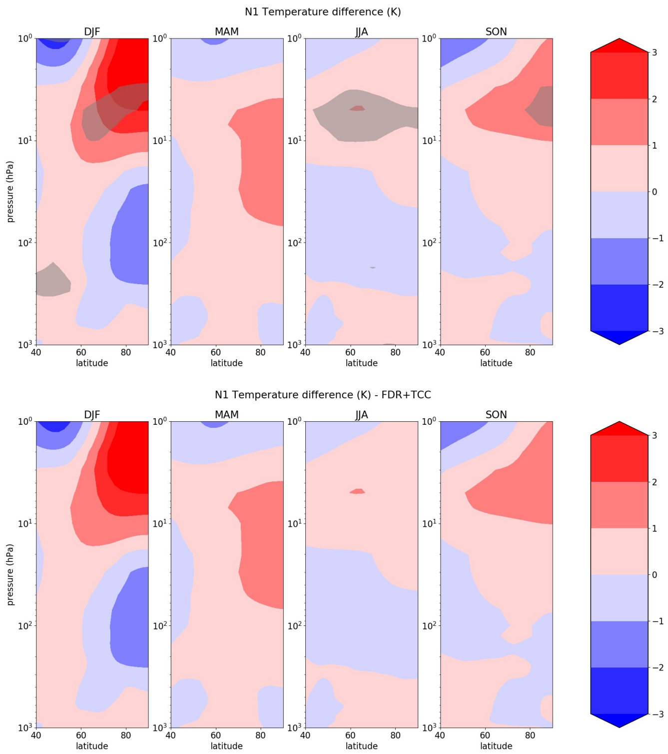

Fig. 1. Zonally averaged temperature difference (High Ap - Low Ap) without corrections (upper panel) and with corrections (lower panel). Gray areas indicate significant differences at 0.05 level.

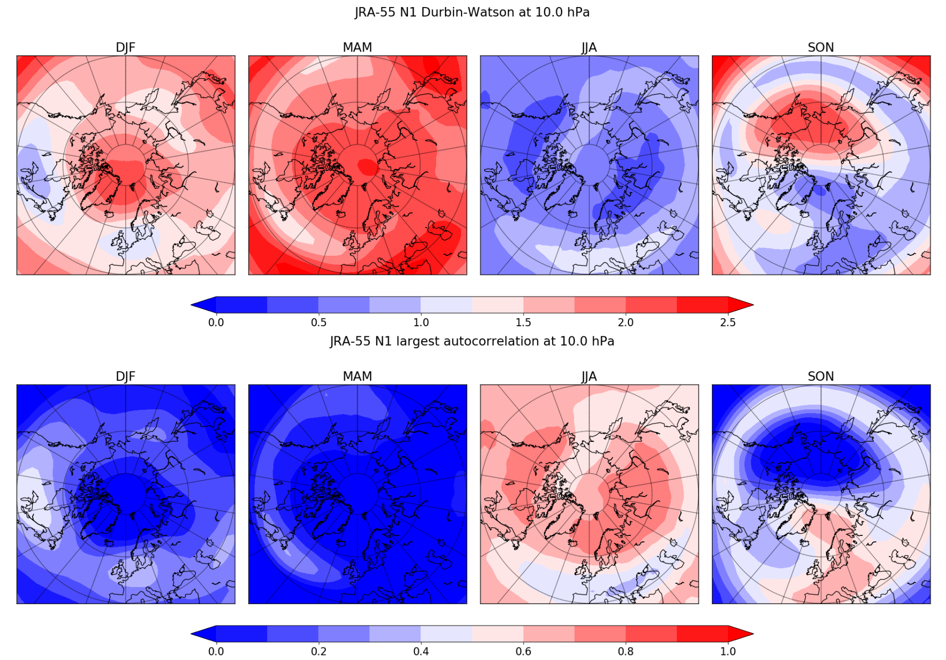

Fig. 2. 10 hPa Temperature - Durbin Watson test result (upper panel) and lag-1 autocorrelation (lower panel) for 10 hPa temperature

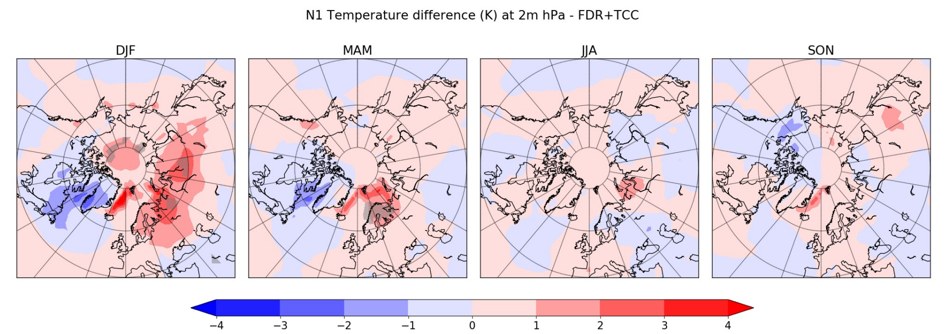

Fig. 3. 2m temperature difference (High Ap - Low Ap) after spatial and temporal autocorrelation temperature. Gray areas indicate significant differences at 0.05 level.