Ann. Geophys. Discuss.,
https://doi.org/10.5194/angeo-2019-156-RC1, 2020

Annales Geophysicae Discussions

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.

Anonymous Referee #1

Received and published: 1 January 2020

In their manuscript, Tartaglione et al. focus on the issue of geomagnetic activity manifestations in local seasonal stratospheric temperatures. Particular attention is paid to the effects of autocorrelations and spatial relations on the formal statistical significance of temperature differences between low/high geomagnetic activity periods. The authors demonstrate that although existence of multiple regions of statistically significant temperature responses is indicated by the basic version of the t-test (i.e., version assuming mutual independence of the temperature values), these disappear when temporal autocorrelations and test multiplicity are considered.

By addressing the subject of statistical testing of multivariate data with notable spatiotemporal correlations, the authors tackle a vital (and arguably still rather underinvestigated) area within the atmospheric research. The text is well written and structured and topically suitable for publication in Annales Geophysicae. Still, there are a few questions and formal issues that the authors should consider addressing when preparing the final version of the manuscript:

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.

(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.

(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.

(C4) A requirement of Gaussianity is mentioned with regard to the t-test (l. 148), 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. 175+), but perhaps this could be mentioned explicitly?

(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.

(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

Minor/technical remarks:

l. 48: “and is thereby” to “and are thereby”

Table 1: “2001” misspelled as “20001”

l. 127: Did Seppälä et al. (2009) really use daily-step data in their analysis?

l. 141: 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)

l. 219: Use of parentheses for rank-ordered p-values is introduced, but parentheses seem to be missing from the relevant formulas in the following text

l. 230: Welch’s variant of the t-test is mentioned (i.e., the form assuming unequal variances of the samples compared), yet t-test employing pooled variance is presented

earlier in the text (eq. (2))

Fig. 3: The green outlines seem to be only partially drawn

l. 264: “point” to “points”

l. 300: extra comma

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