Comment on wcd-2021-68
Anonymous Referee #2

Referee comment on "Differentiating lightning in winter and summer with characteristics of the wind field and mass field" by Deborah Morgenstern et al., Weather Clim. Dynam. Discuss., https://doi.org/10.5194/wcd-2021-68-RC2, 2021

Review of wcd-2021-68
Authors: Deborah Morgenstern et al.
Title: “Differentiating lightning in winter and summer with characteristics of wind-field and mass-field”
Recommendation: Acceptable pending major revisions

Overview

This manuscript explores the meteorological processes leading to lightning in winter and in summer in Northern Germany in an attempt to isolate conditions that differentiate between lightning and no-lightning in summer and in winter. The authors use a pure data driven approach selecting cluster analysis and principal component analysis to group parameter derived from ERA5 reanalysis data into physically meaningful groups representing wind-field-dominated and mass-field dominated lightning conditions.
While this manuscript is generally well written and informative, I have a number of suggestions that follow. Additionally, I have two major concerns that are highlighted below, which will require more thought and effort.

Major Comments

1) Northern Germany was chosen as study area to represent an area in the mid-latitudes and in
flatland to minimize topographical triggering influences on lightning. The selected area is rather small. The results obtained in such a restricted area cannot be transferred to southern Germany nor to the rest of central Europe, which includes the Alpine region. Minimizing topographical triggering influences on lightning is a considerable restriction. Orography often acts as a trigger mechanism in summer as well as in winter. The limitations of the choice of the study area are not stated clearly in the manuscript. What would the results look like for other regions, e.g. southern Germany, the Alpine region or along the Mediterranean coast? Are the findings valid in other regions of (central) Europe? If not, how do they variate?

2) The study is limited to the two seasons winter (December, January and February) and summer (June, July and August). What about spring and autumn? How would the cluster analysis and principal component analysis deviate in these seasons from the presented seasons? Is it really necessary to group the data into seasons? If lightning occurs or not should not be defined by the calendar but rather by the synoptic and atmospheric conditions that lead to the mechanisms of thunderstorm generation and the formation of lightning.

**Minor/Grammatical Comments**

Line 22: “copious amounts of moisture” – What type of moisture is meant here? Moisture near the ground or in the atmosphere? Moisture in form of specific humidity or relative humidity?

Line 54: Figure 1

Line 68: Although a .csv file is provided with a detailed description of the variables used in this study, I suggest to add a table with name, unit and meteorological category of the selected variables.

Line 71: How are equally-sized subsamples formed? Randomly?
Lines 81-92: Can you provide a time series of all lightning that occurs within the selected domain? I assume there are years with higher lightning activities in the respective seasons? How were the “cell-hours” selected? Are they representative for an “average” season?

Lines 110-112: It is not entirely clear to me exactly how this is done. Please add more information or a concrete example.

Line 119: What does “yielding k=5 clusters” mean?

Line 129: Use PCA as this was already introduced in line 120.

Figure 2: Variables derived from ERA5 reanalysis data are not observations. The labeled arrows are very hard to read. Could the readability of the figure be improved?

Figure 3: How would the stacked bar plot of the clusters differ for spring and autumn?

Line 148: Figure 2. Please be consistent with the usage of Figure x or (Fig. x) throughout the manuscript. E.g. use Figure in the text and Fig. in brackets.

Line 188: Within the mid-latitudes, there are many regions where lightning occurrence does not peak in summer months. Please check Taszarek et al. 2019.

Line 274: Is the approach really purely data driven and completely independent of the season?

Line 285: What are “large amounts of CAPE” values?

Line 295: Please check the order of the references throughout the manuscript. The years of the publications should be ascending.
Line 346: What are "substantial amounts of CAPE"?

References: