Climate of the Past Discussions

Clim. Past Discuss., referee comment RC1
https://doi.org/10.5194/cp-2021-52-RC1, 2021

Comment on cp-2021-52

Anonymous Referee #1

Referee comment on "Climate variability and grain production in Scania, 1702–1911" by Martin Karl Skoglund, Clim. Past Discuss., https://doi.org/10.5194/cp-2021-52-RC1, 2021

General comments

This study analyses the relationship between climate variability and grain production in southern Sweden for the 18th and 19th centuries. In the long introductory part, the author provides detailed information about the farming system of the study area, historical background, natural conditions, grain crops, and their varieties. He stressed the importance of crop diversity for stable crop production. It is interesting that reversed relationship of the crop production to temperature compared to other parts of Scandinavia, as well as other parts of Europe, was found in this study.

Historical database of the tithe records is used for the period before 1865 while official statistics on county level were utilized after that year. Data on grain production from the historical database were pre-processed in several steps attempting to solve some biases and sources of uncertainty (normalizing, de-trending). However, not all pre-processing steps are sufficiently explained. For instance, what is the role of the "threshing coefficient". While the reasons for aggregating data are well explained, the application of the cluster analysis (Section 2.2) is rather strange. How it was decided that exactly three clusters are optimal? The final number of clusters was decided subjectively or any objective measure was used? While individual villages are clustered, the map in Figure 3 presents administrative units belonging to different clusters. It would be useful to explain more clearly, whether the BISOS data from the 1865–1911 period were standardized in a similar way as the HDSA data.

The manuscript is not well structured. While the introductory part is rather long, with numerous details, methods are mixed with results in Sections 2 and 3. Section 2.4 is followed by Section 3.1. Some section titles do not correspond to the following paragraphs (e.g. Section 1.4.1). Results of correlation analysis (Section 3.1) are mixed with discussion (e.g. lines 496–504). Results are presented in the form of several correlation matrices (Figs. 9–13). Correlation coefficients between crop production series and temperature/precipitation/drought characteristics were calculated for two different periods, for different clusters and different crops. And these correlations are repeated for

"dry" and "wet" years. The description and interpretation of such results are rather long and not very synoptic. It is very hard to orient in the text and to find any signal in presented correlations. I miss any information which correlations are statistically significant (those in colored boxes?) and on which level (p-values)?

In spite of a relatively long and detailed introductory part, I miss any direct information on harvest dates (or threshing dates?). This date may indicate the time when the grain production of the given year was determined. Harvest dates are mentioned only indirectly in August and September (lines 199 – 201). In this sense, a correlation of a grain harvest with Oct, Nov, and Dec climate of a given year seems to have no meaning. Contrary, it would be more meaningful to correlate the grain harvest of a given year with Dec, Nov, and Oct climate of a previous year. It would be useful especially for crops sown in autumn.

While the analysis of the relationship between grain production and climate is based on simple correlation analysis, a relatively extensive discussion mentions a number of aspects that were not analyzed in this study. For instance, lines 584-585: "... the absence of a climate signal in the spring and autumn months, as well as the last summer month of August to some extent. This could be interpreted as spring or autumn frosts not being a systematic threat …" However, correlation does not mean causality and it would be correct to add some info about the frequency of spring/autumn frosts.

Similarly, lines 630-631: "If conditions were relatively wet or dry in the early summer, the effects from subsequent precipitation and temperatures later in June and especially July would theoretically have been amplified." Such claims should be supported by their own results and / or citations from other similarly focused studies.

Specific comments

- Map of the study area with outlined geography mentioned in the text would be very useful
- Section 2.1 refers to clusters. However, clustering is explained later in Section 2.2.
- Homogenization of the Lund temperature measurements from 1753 is an important by-product of this study. However, this series was also extended further back to 1701 and there is no information about the validation of this earliest part of the "calculated" series. Looking at Figs. 6 and 7, the calculated series (before 1753) seems to have lower variability compared to part of measured temperatures (after 1748) for all seasons. Was the variability of the calculated series adjusted in any way?
- It would be useful to add some information to Table 2 about the length (N) of the period that was used to calculate correlations between the Lund series and the other temperature series. Are there all correlations statistically significant? At that level?
- The problem of homogenization of the early instrumental temperature measurements closely relates to the so called "the early instrumental warm-bias" (see e.g. Böhm al., 2010). It would be useful at least to comment on it. While in the Greater Alpine Region warm bias was found especially in summer, this study found "cold" bias for Lund.
- Table 2 – It is not clear, how the spatial correlation was calculated and why it is listed

on the last row of the table. One would expect that it is a correlation between Lund and neighboring stations (or some spatial temperature field?). In this sense, it would be listed as a separate column, not a row.

- From correlation analysis it follows that explained common variance (r-squared) is mostly negligible. For instance, when the correlation coefficient r=0.3, grain harvest and climate share less than 10% (r^2=0.09) of the common variance. It would be useful to explain more (or even quantify) the role of other factors. Some of them are mentioned in the discussion.

Technical corrections

- Lines 25-25: reference to "Huhtamaa & Helama, 2017b" is mentioned twice in the list
- Line 34-35: correct to (Osvald, 1959; Persson, 2015).
- Line 78: correct to: … century.
- Line 97: "…the suly of winter fodder". Please check. Is it correct?
- Line 109: „starting in 1749/1757" – this is not clear
- Footnote 4: „The share of oats WAS quite low"
- Line 289: please unify: BiSOS or BISOS?
- Line 302. N*ormalized production anomalies is abbreviated as* NPAa, but differently in formula (1) and in the text (line 306)
- Line 347: The data … has subsequently incorporated … Please check
- Line 358: "four clusters" – Should not be "three"?
- Line 382: "…increase in an ascending order …"
- Line 447-8: What is the meaning of „simple" climate variables?
- Line 459: What is the meaning of "…most consistent coefficients"?
- Please check figure and table captions, correct and complete. For instance, there are no „Descriptive statistics" in Table 2, Figure 5 – "…estimated loess" is not clear. The loess function is used here as a low-pass filter. Figure 7 – correct the caption – figure relates to DJF and SON seasons. Figures 9 – 13 – please add information about the statistical significance of the correlation coefficients.