CC again as RC

Christian Pfister (Referee)

Referee comment on "Quantifying and reducing researcher subjectivity in the generation of climate indices from documentary sources" by George C. D. Adamson et al., Clim. Past Discuss., https://doi.org/10.5194/cp-2021-167-RC1, 2022

Review of Adamson et al: Quantifying and reducing researcher subjectivity in the generation of climate indices from documentary sources

The study aims exploring and quantifying the degree of error between researchers assigning ordinal-scale indices to a historical documentary dataset. Two teams of raters were asked to produce a five-category annual rainfall index series for a dataset consisting of transcribed narrative descriptions of meteorological variability for 11 rain-years' in nineteenth-century Lesotho. The authors conclude that variability between researchers should be considered minimal where index-based climate reconstructions are generated by trained historical climatologists working in groups of two or more.

The study should be accepted with small changes indicated below:

- The different preconditions in the derivation of temperature and precipitation indices should be worked out more clearly :

Lines 49-56

"This approach has been adapted for regions with less rich documentary evidence", **"The Pfister is approach is mainly tailored to reconstructing temperatures for regions with rich documentary evidence and long series of homogenized instrumental measurements (e.g. Pfister, Wanner 2021). In such cases proxy information often allows estimating temperatures for specific months or seasons by using the calibration verification approach (e.g. Dobrovolný 2010). In such cases, the potential bias in classification is very small for trained historical climatologists, as the narrative record and the proxy need to be consistent and meteorologically meaningful.** This approach has been adapted for regions with less rich documentary evidence, or a seasonal skew to the available climate descriptions, through a reduction in the number of index categories (e.g. to five or three classes) and/or the temporal resolution of the reconstruction (to seasonal or annual).

The situation is different for classifying precipitation. Proxy-data such as information on floods and droughts or the number of rain-days may hardly be calibrated, as precipitation is rather small-scaled in comparison to temperatures and because long homogenized instrumental series of precipitation are quasi non-existent. The study by Dobrovolný et al.(2015), which is perhaps the most sophisticated approach of this kind in Europe, only indicated acceptable reconstruction skill for seasonal precipitation indices in JJA and annual values.

- It should be worked out more clearly, also in the abstract, why estimates of historical precipitation conditions in a country of the Global South are significant for the present situation.
- a map should be included showing the location of Lesotho in southern Africa.
- an example of a source illustrating the nature of the narratives should be included.

Suggestion for small changes:

Line 45: "Under the Pfister method, indices are normally" might be replaced by **"The Pfister Indices, as Mauelshagen (2010) named them"**, are normally generated,

Lines 49-51: "relevant phenomena (e.g. the timing and duration of snowfall, or various plant-phenological indicators) "might be replaced by . "regionally relevant **proxy data (e.g. plant-phenological observations, the duration of snow-cover and the freezing of water bodies)**…

Additional references:
Mauelshagen, F. (2010), Klimageschichte der Neuzeit, 1500–1900. Darmstadt: Wissenschaftliche. Buchgesellschaft, 2010
Pfister C. and Wanner H. (2021). Climate and Society in Europe,