

“Climate indices in historical climate reconstructions: A global state-of-the-art”

Changes to manuscript in response to review comments – second revision

Response to Referee #1

I appreciate the work that the authors have taken in revising the manuscript. Some of my comments have been taken into consideration but it is clear that we do not share the same view of what a critical review is. In their reply they say that ‘they embed critical analysis throughout the manuscript’, but sections 1-8 are still purely descriptive of the current state of the art, which, as told in my previous review is a highly valuable effort. The new version of section 9.2 has some more analytical character but limited to the treatment of the uncertainty. I think that a stronger comparative analysis among the different approaches would lead to a more convincing paper, but I do not want to go further in this controversy.

Response – Many thanks for these comments. The main focus of the manuscript is – as described in the introduction – to document previous work that has used an index approach to reconstruct historical meteorological entities. We politely disagree that the paper is ‘purely descriptive’ and can point to numerous examples of critical review and comparative analysis. The most obvious ones are our discussion of the varying number of index points used in European reconstructions (section 2.2), our comments on the use of the ‘0 index’ value across continents (see text throughout section 8), the comparisons of the results of overlapping reconstructions for Africa (section 8.3) and our revised discussion of confidence and uncertainty (section 9.2).

We have added a caveat to the last paragraph of the introduction to frame the review for the reader: *“...the emphasis of the article is on the documentation of studies that have used an index approach to climate reconstruction, with critical review and comparison where appropriate. The number of instances where comparative analysis is possible is necessarily restricted by the limited number of studies that have undertaken either different approaches to index development for the same location or identical approaches for different regions.”*

The guidelines have been partly modified following my comments, but I still have problems here.

Guideline 8. I do not agree how it is formulated. I think that a general statement advising to sum index series is misleading. Take the case of wet indices. What would happen if you add -3 (very dry period) and 3 (very wet), you will get a 0, indicating a misleading normal value, while what you actually got was two extremes within the same period. The index aggregation should be done with careful assessment, so this guideline needs to be reformulated. This comment also applies to the new text inserted in section 8 after table 1.

Response – The reviewer is correct that, when summing the scores for individual months/seasons to produce seasonal/annual totals, negative and positive values cancel one another out. This is perfectly acceptable and is the main way in which years with a ‘0 index’ arise. We appreciate, however, that the process may result in a loss of information.

We have amended the paragraph below Table 1 in section 8.1 to read: *“Once monthly index values have been generated, these are then summed to produce seasonal or annual classifications where required. Three-month seasonal values can, as a result, fluctuate from -9 to +9 and annual values from -36 to +36 (see Pfister, 1984). The process of summation may result in positive index values for relatively warmer/wetter months during the year being cancelled out by negative index values for relatively colder/drier months. For example, a year containing a run of extremely dry months followed by a run of extremely wet months may produce a summed index value close to zero – even though the year includes two periods of ‘extreme’ climate. Careful assessment is therefore required when reporting summed indices to avoid any loss of information, particularly concerning extreme events. The approach used by Nicholson et al. (2012a) for African precipitation series may be helpful here, where individual years were flagged if documentary*

sources suggested wetter and drier extremes across the year that differed by more than two index classes.”

Guideline 8 has also been amended to read: “It is advisable to sum-up index series – either in time (i.e. from monthly to seasonal or annual) or in space (i.e. by combining several index series from a climatologically homogeneous region). This approach may well approximate index series to natural climate variability. *Careful assessment is needed, however, to avoid any loss of information during the process of summation, particularly for extreme events (see section 8.1).*”

Guideline 10. The resolution is not the problem when calibrating indices against proxies. The problem is that proxies are an indirect indication of the climate, with their own uncertainties and caveats. They are related through a transfer function with the climate variable and they cannot be taken as the ‘truth’ for calibration. You can use them for testing and comparing but should not be considered for calibrating, since the index might have a more direct relationship with the climate variable than the proxy itself.

Response – Thank you for this. We have modified the sentence to read: “Where instrumental data are not available, overlaps with independent high-resolution palaeoclimate records may be *useful for comparison and testing, noting that palaeoclimate records may have their own biases.*”

Other questions

The statement (lines 713-714 in the track changes manuscript) about precipitation following a gaussian distribution should be corrected. Different from temperature, the precipitation follows different distributions depending on the considered scale, from Gamma (for short periods) to gaussian. The sentence in the text should be clarified.

Response – The reviewer is correct. Precipitation does not follow a Gaussian distribution at short time scales. However, when dealing with monthly time scales (which Pfister was), a broadly Gaussian distribution can be assumed. We have clarified the text to read: “In the development of his seven-point scale, Pfister assumed that *monthly* temperature and precipitation followed a Gaussian distribution”.

I do not understand the sentence in lines 924-926. “However, even where a period of overlap is lacking, indices from documentary sources can still be used to test reconstructions from proxy data or reconstruction, modelling results and observations”. I do not understand the last part of the sentence.

Response – We have revised this sentence for clarity to read: “However, even where a period of overlap is lacking, indices from documentary sources can still be used to *cross-check* reconstructions from proxy data (...) or modelling results and observations”.

I am wondering why the well-known CLIWOC figure 10 is not referenced to the original papers.

Response - We did not reference the original paper for Figure 10 because the image is drafted from data in the open source variant of the CLIWOC database stored at [historicalclimatology.com](https://www.historicalclimatology.com). We have updated the figure caption to clarify this: “Figure 10: Plot of the position of all ships’ logbook entries in the CLIWOC database (Degroot and Ottens, 2020). The map is derived from the open source variant of the CLIWOC database (García-Herrera et al., 2005b) held at <https://www.historicalclimatology.com>.”

One of the longest series derived from documentary sources is the Liu et al 1000 year series of typhoons in S China. This should be included in the text.

Kam-biu Liu, Caiming Shen & Kin-sheun Louie (2001) A 1,000-Year History of Typhoon Landfalls in Guangdong, Southern China, Reconstructed from Chinese Historical Documentary Records,

Response - We did not include this paper in the review as it concerns historical typhoon occurrence/frequency rather than an index-based reconstruction (e.g. of typhoon severity).

Response to Referee #2

I recommend to do only several minor changes in the final version of the manuscript (see below). I believe that the whole article and especially the final guideline will represent a significant contribution to the development of historical climatology.

Response – Many thanks for this kind comment.

Suggested minor changes:

Line 712: Rather there should be „...calibration and verification... “. However, as the whole paragraph is dealing with the process of index development, the last sentence may be omitted.

Response – we have revised the text to read ‘...calibration and verification...’ We have retained the last sentence as it is referring to the next section of the document rather than the following paragraph.

Line 720: “... from the mean of the reference period)”

Response – thank you for spotting this – corrected as suggested.

Lines 933–934: please consider formulation that is more correct: “During verification, index values calibrated to physical units (e.g. temperature degrees or precipitation amount) are compared with the instrumental data ...”

Response – thank you – the text has been changed as suggested.

Line 939: Please consider using e.g. “...not stable through time...” instead of “non-stationary”. Several types of “stationarity” (and non-stationarity) may be defined strictly in statistical sense.

Response – thank you – text has been amended to read: ‘...may not be stable through time...’

Line 1024–1025: There should be “... at the risk of incorporating non-homogeneities.”

Response – thank you – corrected as suggested.

Response to Referee #3

The manuscript has improved a lot in the review process. Nevertheless, there are some important questions than remain without solution and I cannot recommend its publication in its current form. One major point is that the authors do not accept the lack of critical analysis in the in the manuscript suggested by RC2 and SC1. In my opinion, to infer the guidelines proposed at the end of the manuscript by the previous sections continues to be impossible.

Response – This is a difficult comment to address, as the reviewer provides no specific instances to back up his or her viewpoint. However, please see our response to the opening comment by Reviewer #1, as this addresses Reviewer #3’s concerns about the level of critical analysis within the manuscript.

The authors are excluding of the review some indexing methods and documentary sources (see paragraph 197-203 in the manuscript, page 12 of the reply and comments below). For this reason, it is necessary to clarify on which documentary sources and type of reconstructions is focused the review. In the pag.12 of the reply letter the authors said that the narrative sources are the key focus “Space precludes a full coverage of these and comparable indices relating to Europe, so we will emphasise that our key focus is on narrative sources”. In this case I suggest including this concept clearly in the title and abstract. Moreover a clear definition of “narrative source” in the introduction is required, and modifies the manuscript according with this definition. I understand that not all the cited sources in the introduction can be considered narrative sources “Information sources include, but are not limited to, annals, chronicles, inscriptions, letters, diaries/journals (including weather diaries), newspapers, financial, legal and administrative documents, ships’ logbooks, literature, poems, songs, paintings and pictographic and epigraphic records” and many indices cited in the text are based on both narrative and no narrative sources.

I have the feeling that the authors want to focus the manuscript in situations in which the researcher has many different documentary sources with different origins and with highly variable climate information. In this situation, probably the categorized index +X/-X is a good option. If this is the idea, this must be crystal clear in the manuscript, starting by the title because currently is too general. In this case probably has no sense to include in the review the works focused on weather diaries, logbooks, phenology or chapter acts among others, because in these cases researchers use only one type of documentary sources with very specific climate information. Another option would be to divide the review in works that use many different documentary sources with high variable climate information to reconstruct one series and the works that use only one kind of documentary source with specific climate information.

Response – This comment combines several concerns, which we respond to thematically.

1. Narrative evidence/narrative sources

The reviewer requests that we provide clearer definitions of ‘narrative sources’ and ‘narrative evidence’. Although the term is widely used, defining a ‘narrative source’ is not easy. Documentary sources fall along a spectrum from those containing purely quantitative data to those that include purely narrative evidence.

Rather than attempt a definition, we have checked every mention of ‘narrative source’ in the manuscript and changed it to ‘sources containing narrative evidence’, ‘documentary sources’ or ‘narrative evidence’ as appropriate. We have also made minor edits to the abstract and opening paragraphs of the manuscript to provide greater clarity. Note that (following Brönnimann et al., 2018) we already identify that narrative evidence is concerned with “descriptions of short-term atmospheric processes and their impacts on environments and societies” (lines 35-36) and echo this in line 41 where we state that narrative descriptions contain “local observations of short-term atmospheric processes and their impacts”.

2. Exclusion of indexing methods and documentary sources

The reviewer suggests that we are “excluding of the review some indexing methods and documentary sources”. This comment appears to be related to his/her view of our treatment of historical records of rogation ceremonies. We have hopefully addressed this concern in our response to his/her final substantive comment (see final page of this response). As we explain in the manuscript, we are selective only in the text dealing with Europe, where there is such a large volume of index-based studies that a book would be required to do it justice. Even here, we are very clear that the review only excludes studies that do not include original published series based on primary sources and/or those that do not reconstruct meteorological entities (lines 93-96).

3. Single versus multiple sources

The reviewer suggests that we “want to focus the manuscript in situations in which the researcher has many different documentary sources with different origins and with highly variable climate

information”. This isn’t quite correct. In the paper, we consider studies where authors have used both single source types (e.g. Brázdil and Kotyza [1995, 2000], Fernández-Fernández et al. [2014]) and multiple source types. We also consider studies that include narrative evidence from weather diaries and from phenological records. It should be remembered that these types of source may contain descriptive accounts of meteorological phenomena as well as quantitative data. In all of these cases it is perfectly feasible to develop categorised indices in the format +X/-X. With this in mind, we would prefer to keep the manuscript organised on a continent-by-continent basis.

We agree, however, that some source types are not suited to the development of ‘+X/-X’ style indices. With this in mind, we have edited the second paragraph in section 10.2, as follows: “Having a standard approach to index-based climate reconstruction would clearly have its benefits. However, we recognise that a ‘one size fits all’ approach is *neither* appropriate for all climate phenomena *nor for all source types*. The reconstruction of historical wind patterns over the oceans from ships’ logbooks *and the identification of precipitation variability through the analysis of descriptions of rogation ceremonies*, for example, already *have* well-developed methodologies and protocols.”

We have also amended the text immediately preceding the guidelines: “The guidelines are of greatest relevance to index-based reconstructions of temperature and precipitation *from multiple source types* but also have resonance for other climate phenomena (e.g. winter severity) *and for many single source types* (e.g. *annals, chronicles, letters, diaries/journals, newspapers*).”

In my opinion this paragraph:

line 197-203: Index series based on historical records of rogation ceremonies – closely linked to precipitation (or a lack thereof) – warrant separate discussion. This source type is particularly valuable for western Mediterranean regions (e.g. Álvarez Vázquez, 1986; Martín-Vide and Vallvé, 1995; Barriendos, 1997; Piervitali and Colacino, 2001; Domínguez-Castro et al., 2008; Barriendos, 2010; Domínguez-Castro et al., 2010; Garnier, 2010; Domínguez-Castro et al., 2012b; Fragoso et al., 2018; Tejedor et al., 2018; Gil-Guirado et al., 2019; Bravo-Paredes et al., 2020). However, as most studies base their indices on the type or cost of ceremonies – or the space within individual documents devoted to describing each ceremony – rather than a meteorological entity, we do not go into further detail.

has some conceptual errors and it is incorrect. The articles cited only have in common that they use rogation as a climate proxy. Some articles generate a precipitation index using pro pluvia and pro serenitate rogations i.e. Álvarez Vázquez, 1986; Martín-Vide and Vallvé, 1995; Barriendos, 1997; Barriendo 2010. Other as Fragoso et al. (2018) work with different climate information, not only with rogation and generate a monthly precipitation index (-1, 0, +1) and a drought record. Gil-Guirado et al. (2019) provides a drought index mainly based in pro pluvia ceremonies and an extreme rainfall index based in some pro serenitate rogations and much other climate information. Bravo-Paredes et al. (2019) analyse the use of pro pluvial rogations in Extremadura as winter NAO proxy. Finally, Piervitali and Colacino (2001), Domínguez-Castro et al. (2008, 2010, 2012b), Garnier (2010) and Tejedor et al. (2018) produce drought index from pro pluvia ceremonies information with different methodologies of indexation. For this reason, in my opinion some articles are correctly cited in section 2.3 but others have no sense in there and must be cited in section 2.5. In addition probably is better to cite the final version of Tejedor et al., 2018/19 accepted in climate of the past (Clim. Past, 15, 1647–1664, 2019, <https://doi.org/10.5194/cp-15-1647-2019>) instead the Climate of the Past Discussion version.

As I am mentioning, all these works provide indices of “meteorological entity” i.e. precipitation, drought or extreme rainfall events. Moreover some of these works provide original indexation methods far from the (-X,+X) continually cited in this review. In my opinion these original methodologies deserve to be commented and analysed in this review. Dominguez-Castro et al (2008) and Gil-Girado et al (2016) are methodological works, in which different indexing methods are compared. In my opinion this kind of works are necessary to know which methodologies of indexation are better, and not only try to accept and repeat the legacy of previous studies as the authors recommend in the point 4 of the guidelines.

Response – We thank the reviewer for this helpful detail. We have edited and expanded the text in the final paragraph of section 2.3 such that it provides much more specific information about rogation as a precipitation proxy:

“Index series based on historical records of religious rogation ceremonies warrant separate discussion. Rogations are liturgical acts conducted to request either rainfall during a drought (termed pro-pluvia rogations) or an end to excessive or persistent precipitation (pro-serenitate rogations), and were used as an institutional mechanism to address social stress in response to such meteorological extremes (see Martín-Vide and Barriendos, 1995; Barriendos, 2005; Tejedor et al., 2019). Analyses of the occurrence and nature of rogation ceremonies have proven particularly valuable for western Mediterranean regions (most notably the Iberia Peninsula), where they have been used to create precipitation indices spanning the 16th to 19th centuries (e.g. Álvarez Vázquez, 1986; Martín-Vide and Vallvé, 1995; Barriendos, 1997, 2010; Gil-Guirado et al., 2019). In some cases, information about rogation ceremonies has been combined with climate-related narrative evidence to generate precipitation series (e.g. Fragoso et al., 2018). Useful evaluations of different indexing methods are provided by Domínguez-Castro et al. (2008) and Gil-Guirado et al. (2016). For a discussion of the use of rogation ceremonies as a proxy for drought see section 2.5, and for examples of rogation-based reconstructions in South America see section 5.”

We have added the following paragraph to section 2.5 on drought indices:

“Drought indices have also been derived for the Western Mediterranean using records of rogation ceremonies, with specific methodologies developed to estimate the length, severity and continuity of drought episodes (see Domínguez-Castro et al., 2008). A number of studies have used evidence of pro-pluvia ceremonies (see section 2.3) as a drought proxy (Piervitali and Colacino, 2001; Domínguez-Castro et al., 2008; Domínguez-Castro et al., 2010; Garnier, 2010; Domínguez-Castro et al., 2012b; Tejedor et al., 2019), sometimes in combination with other narrative evidence (e.g. Fragoso et al., 2018; Gil-Guirado et al., 2019). Readers are referred to Brázdil et al. (2018) for a detailed discussion of the different types of drought indices.”

And the following sentence to section 2.6:

“Pro-pluvia rogation ceremonies have been analysed as a proxy for the winter North Atlantic Oscillation between 1824 and 1931 CE in the Extremadura region of Spain (Bravo-Paredes et al., 2020).”

Finally, we have updated the reference for Tejedor et al. to the accepted 2019 version in *Climate of the Past*.