

Dear Prof. David Nash,

Thank you very much for your kind, considered, and very constructive comments. We have made several changes based on your queries and recommendations. Apart from including them in a revised draft, we have detailed our responses below.

Best,

The Authors

General comments

This is a welcome paper addressing the thorny issue of how best to integrate historical and meteorological evidence to reconstruct climates of the past. It is clearly structured, generally very well written and makes persuasive arguments. However, as you will see from my specific comments below, it includes some statements that are problematic. These mainly relate to the way that the text falls into the “documentary evidence unreliable, instrumental data reliable” trope common to many climatological studies.

Agreed. We have made several edits to revise the tone of the work so that the documentary evidence is given more of its due. Our intention was certainly not to diminish the utility and importance of the documentary-derived data – our time-series rely significantly on it, after all. Notable edits are on lines: 51-2, 120-4, 180-2.

There are also some methodological issues that need addressing. These include the need for a clearer description of the way in which documentary index classes are derived.

Descriptions added to Table 1 (line 197).

Explanation is also needed as to why – given that the goal of the study is to produce a time series that is “interoperable with Nicholson et al. (2012)” (line 167) – the method used to convert modelled rainfall levels from 20CR and GCMs is not the same as that used by Nicholson et al. (2012).

This error stems from Philip Gooding (Author 1) misinterpreting what ‘interoperable’ means – apologies. Now that he has informed himself, we can state that the aim was not to make the dataset interoperable with Nicholson et al. (2012). Instead, we think our outputs, which use a 7-point scale, may be

“broadly comparable” (edit made) with the Nicholson et al. (2012) dataset because they use the same scale, even if they are not entirely interoperable because we use different sources and methods. The choice to use the 7-point scale was made because the data was in many places granular and detailed enough that we could be this precise (instead of limiting ourselves to a 3-point or 5-point scale). Again, edit made to reflect this. We think this edit also helps to respond to the reviewer’s first point (lines: 177-82)

I’m very surprised that the authors do not cite the recent papers by Nash et al. (2018) and Mutua & Runguma (2020), which present 19th century documentary climate series for Malawi and Kenya respectively. I would have thought that these are essential for comparison with the results presented for Tanzania.

Thank you for these references. We’ve read them and made a number of changes that incorporates evidence from them. We’ve also added evidence from rain gauges in Kampala (1879, 1881-1886), Mombasa (1875-1881), and Zanzibar (1874-1881), which are referred to in the data for Nicholson et al. (2012). These help to situate the findings with data already gathered from the wider region. See esp. lines: 93-5, 123-4, 469-79.

Mutua, T.M. and Runguma, S.N. (2020) Documentary driven chronologies of rainfall variability for Kenya, 1845–1976, *Journal of Climatology and Weather Forecasting*, 8, 255, available at: <https://www.longdom.org/open-access/documentary-driven-chronologies-of-rainfall-variability-for-kenya-18451976.pdf>

Nash, D.J. et al. (2018) Rainfall variability over Malawi during the late 19th century, *International Journal of Climatology*, 38 (Suppl. 1), e629–e642.

Specific comments

Lines 16-17, 60 and 507 – I’m not sure the wording “...a more scientifically grounded interpretation of documentary materials...” in the abstract and main text is ideal. This makes a value judgement about the validity of documentary evidence. Maybe ‘climatologically grounded’ rather than ‘more scientifically’ grounded would be better.

This is a good point. We have changed it to, ‘an interpretation of documentary materials that is grounded in both the humanities and natural sciences.’ See lines: 16-17, 63-4, 560.

Line 38 (and throughout) – best to avoid the use of terms derived from ‘East Africa’ as this is a colonial construct. Academics that I have worked with from the region tend to prefer the term ‘eastern Africa’.

This is true, and something Philip Gooding (Author 1) addressed in his recent monograph (*On the Frontiers of the Indian Ocean World: A History of Lake Tanganyika, c.1830-1890*, Cambridge, 2022, p. xii). Unfortunately, ‘eastern Africa’ doesn’t really work either, as the term can be (and has been) applied to anywhere in the eastern half of the continent, from Cairo in the North to the Cape in the South. Thus, it lacks specificity. The better term is ‘equatorial eastern Africa,’ but this is perhaps overly long and its usage may hinder readability. We also noted that several climatological studies (e.g. Alin and Cohen 2003; Bessems et al., 2008; Hastenrath 2001; Nicholson 2015; Nicholson and Yin 2001; Verschuren et al. 2000) use ‘East Africa’, and so we thought we were conforming to established practice. But, in the interest of challenging colonial constructs, perhaps this practice needs to change. Thus, we have changed all references to East Africa to equatorial eastern Africa. However, if there are objections related to readability, we would rather revert to East Africa (which is usually understood to refer to Tanzania, Kenya, Uganda, Burundi, Rwanda, and the eastern DRC) than to adopt ‘eastern Africa,’ because of the latter’s lack of specificity.

Line 45 – I don’t think Endfield & Nash (2002) used a term as strong as ‘distorted’ to describe European perceptions (and hence descriptions) of African climate. Rather, the descriptions made by Europeans were often framed relative to their ‘home’ climate (particularly during their early years of residence in Africa), so tend to over-emphasise drier conditions. As noted in section 2, they may also be shaped by imperial knowledge-making.

- We’ve clarified this. True – their understandings of climate on its own were not necessarily ‘distorted.’ However, the role that they saw climate playing in the regions they reported certainly was i.e. they conceptualisation was rooted in enviro-climatic determinism. We’ve made an edit to reflect this and taken out the reference to Endfield and Nash 2002, which does not deal with this issue as much as the other cited texts. See lines: 46-7.

Lines 46-48 – this is a very strong statement – are you sure that absolutely no records made by Europeans describing climate in Tanzania exist between 1861 and 1868? I find that very hard to believe.

This is actually correct, although we’ve made an edit to clarify. No Europeans entered the region between Speke/Grant’s expedition (which ended in 1861) and

David Livingstone's arrival in Ujiji in 1869. The clarification we've made is to point out that there were no first-hand reports between these dates (this is similar to inland Kenya, as Matua and Runguma (2020) point out as well). There may have been the odd second-hand report made by administrators in Zanzibar, who may have consulted with Omani/African traders and/or porters who had returned from inland regions, but we have not made an effort to find them, if they even exist. If they do, they will not provide the kind of granular data that the firsthand reports that we have consulted contain. Line 49.

Lines 114-127 – this paragraph makes some valuable points. However, it paints the rather sweeping picture that all explorer and missionary descriptions of weather and climate were shaped by imperial agendas and are therefore unreliable. Some descriptions might well be 'highly subjective' – especially the broad overviews of climatic conditions in the more general explorer monographs – but other accounts of specific weather events and related phenomena (e.g. delays to the start of the rainy season, counts of rainy days, descriptions of flood events, descriptions of pasture conditions etc) will likely be reliable. I suggest that this paragraph be tweaked to provide greater nuance.

Yes, the paragraph, as it stood, skewed overly negative. Thus, we have added 'positive' aspects of the documentary material as well: "Descriptions of rainy days, flood events, pasture conditions, and harvests, which are all to varying degrees related to climatic conditions, are abundant and can be used to make time-series of rainfall variability (Nash et al., 2018; Matua and Runguma 2020)." (Lines 123-4). We've kept the references to imperial knowledge-making, however. We feel it important to acknowledge how and why these documents exist, as well their utility to scholars in the present (the latter being much more well-known).

Lines 128-129 – the emphasis on describing extreme conditions is not unique to African documentary evidence and is well documented in historical climatology studies around the world – have a look at some of the excellent reviews by Christian Pfister or Rudolf Brazdil for further details and cite relevant methodological sources.

Thank you. Added references to Pfister 1995, Brazdil 2000, and Brazdil et al. 2005, as well as a call out to this wider research in the opening line of the paragraph: "Given this historical background (and in line with other documentary sources for other regions and time periods), it is probably unsurprising that Europeans commented on climatic conditions more when they were extreme, such as in instances of severe drought or floods, than during months/seasons/years of regular rainfall." Line: 135-8.

Lines 130-139 – these kinds of uncertainty surrounding ‘climatically indirect’ indicators of climate variability are routinely dealt with in historical climatology studies and there is a wide literature on this. Again, have a look at some of the reviews by Christian Pfister or Rudolf Brazdil for details. These include explicit guidance on how to handle an ‘absence of discussion’.

Added qualifier: “Absence of discussion about climatic conditions may be indicative of regular rainfall, *especially if there are no or few reports of disruptions to phenomena that are regularly affected by rainfall extremes, such as harvests and travel.* Nevertheless, such an assumption necessarily comes with a degree of uncertainty.” Reference made to Pfister 1995; Pfister et al., 2018. Lines: 146-9.

Lines 144-148 – this is a very long sentence – suggest you fragment.

Agreed. Edit made.

Lines 155-157 – on a more pragmatic note, it is also very likely that they were interested in weather conditions as they relied upon them to grow their own food.

This is implied on line 130 of the original submission (line 139 of the latest version), in which it is stated that droughts and floods could have an adverse impact “on Europeans’ and surrounding societies’ everyday lives...” Thus we’ve made no edit here.

Lines 169-178 – I’m slightly unclear over the methodology used here. Are you following directly the methodology used by Nicholson et al. (2012) whereby individual pieces of narrative evidence (i.e. individual quotes) are read and graded from 1-7 (and then averaged to give an annual picture), or the approach used in most other documentary-based climate reconstructions around the world where collections of quotes from specific months or seasons are read together and given a collective grade? This is important because, as Nash et al. (2021 – section 8.3) have discussed, the Nicholson method tends to lead to an over-representation of drier conditions in the resulting reconstruction. Figs 2-4 seem to suggest some sort of hybrid, which could be problematic if the authors are aiming to replicate Nicholson’s approach as they suggest.

We read and graded pieces of narrative evidence and used the archive to guide us on the time-period to which it referred. For example, a report from Tabora made in January stating that the rainy season had yet to begin would provide data for November-January. Alternatively, a similar report from March, would provide data for November-March. We’ve added sentences to clarify this. We are

not trying to replicate Nicholson et al 2012 – see above for explanation. Edits made on lines: 178-85.

Lines 199-203 – I would appreciate a little more explanation over the way in which individual diary entries are incorporated into Figs 2-4, particularly where they are merged with results from quotations in letters that could refer to conditions over periods of longer than a single day. In effect, you appear to be giving equal weight to (for example) a single daily diary entry describing drought and a letter documenting dry conditions that could span weeks or months. If you are following the Nicholson method described above this could lead to an over-representation of particular conditions, especially if these are isolated quotes from a personal diary rather than a ‘weather diary’ with daily weather-related entries.

Sentences added to clarify this, using the example of the one letter and 12 diary entries. See lines: 202-12.

Line 225 – there are many reasons for famine, not simply climatic. Do you have any contextual data from missionary sources that might explain the causes?

Yes. Added a sentence to reflect this with citations to missionary reports. We have also added a reference to Rockel 2022, which deals with this drought/famine in late-nineteenth century equatorial eastern Africa in depth. Lines: 247-53.

Lines 319-321 – this sounds like the approach used in the majority of historical climatology studies based on the European tradition, where monthly indices are summed and averaged. There is nothing new here methodologically, so you would do well to cite related sources – see section 8.1 in Nash et al. (2021) for more detail.

This is correct. Citations (Nash et al., 2021; Pfister et al., 2018) have been added. Lines: 352.

Lines 337-347 – I’m intrigued to know why you have adopted this approach when Nicholson has published her method for converting rain gauge data into 7-point index values based on standard deviations from the long-term mean (see section 8.3 in Nash et al. 2021 for a summary). If, as suggested earlier by the authors, they are trying to make results that are interoperable with those of Nicholson, then surely the same method needs to be used in this study?

It has now been clarified earlier in the paper that this study is not made to be exactly comparable or interoperable with Nicholson et al. 2012 – see above. We use our methodology (and not that of Nicholson et al. 2012) because it allows

conversion of the reanalysis and GCM data to the same scale as the documentary data, enabling direct comparisons and integration, whilst also maintaining the temporal resolution of both reanalysis and GCM datasets.” (Lines 345-6 in first submission; lines 375-6 in revised submission).

Lines 513-517 – these sentences again oversimplify the apparent subjectivity of European observers. If you are going to make statements that European observers “regularly misunderstood the climatic and environmental contexts they reported on”, then you need supporting evidence. I would suggest softening of these two sentences. There is as much evidence in the literature supporting the idea that European observers provided reliable eye-witness testimonies of climatic conditions in Africa as there is that their observations were unreliable.

Revised to critique historians’ interpretations of documentary materials more than the materials themselves: “A recurring theme is that such reports may often be more indicative of when droughts began to adversely affect the societies within which the Europeans lived or visited, not just when the drought set in. Thus, historical interpretations of, for example, historical famines, may be better understood as the result of an unfolding of longer-term climatic, environmental, and societal factors, rather than the onset of a sudden disaster, as permeates much of the historiography (cf. Rockel 2022; Mutua and Runguma 2020).” Lines: 564-7.

Lines 522-525 – in light of the work by Mutua and Runguma (2020), this sentence requires revision.

Agreed, change made: “Thus, direct evidence from the region in the form of historical documents, hitherto not incorporated into global climate reconstructions, allow the models to be deployed...” Lines: 575-7.