

Reply to RC1

Overview

This manuscript provides an important new reconstruction of winter temperature for southwestern China as well as an important demonstration of, and update to, the Bayesian historical climatology approach introduced in Camenisch et al (2022). The new methods and results of this manuscript look solid. Nevertheless, the presentation needs some improvements. I was already familiar with the Bayesian approach before I read this manuscript, yet I still found it difficult to understand some of its methods and results. I suspect the current manuscript would be confusing or challenging for many readers. Moreover, the manuscript could do a better job presenting both advantages and challenges of the Bayesian approach. Therefore, I recommend publication after revisions. The following review highlights places in the manuscript that need particular attention.

Response: Thank you for the comments and helpful suggestions. We will carefully review and revise the manuscript to improve the overall readability, the use of terminology, and the clarity of methodology. In particular, we will provide a more detailed explanation of the Bayesian approach in the revised manuscript follow a flowchart, and we will add a discussion of the comparison between Bayesian approach with other approaches. Below, we provide point-by-point responses to your specific comments.

Specific Comments

1. Line 22: The qualification “data-sparse” is relative. It would be better to explain that there is abundant data for eastern China but relatively less for southwestern China.

Response: Thank you for the comment. The “data-sparse” of the study area used in the manuscript is relative to eastern China. We have revised the sentence accordingly.

2. L35: Should this be high spatial *resolution*? In most parts of the world, the spatial *coverage* of documentary data is usually not very good, with most observations concentrated in cities or a few rural locations. Is this different in SW China?

Response: Thank you for the comment and for pointing this out. It is true that, in many regions of the world, documentary data have more limited spatial coverage than some natural proxies such as tree rings. However, the situation is different in China, especially for winter temperature reconstructions. In eastern China, documentary records generally provide broader spatial coverage than tree-ring data. In southwestern China, although tree-ring records are available, they are mainly concentrated in the western areas near the Tibetan Plateau and the Hengduan Mountains (Shi, 2017; Li, 2021). Documentary records are more widely distributed across the provinces included in this study, despite their limited continuity. We agree that the advantages of documentary data lie in their accurate dating and rich climatic information

content. Therefore, we will revise this sentence to emphasize these two strengths more clearly.

Reference:

Shi, S., Li, J., Shi, J., Zhao, Y., Huang, G.: Three centuries of winter temperature change on the southeastern Tibetan Plateau and its relationship with the Atlantic Multidecadal Oscillation, *Climate Dynamics*, 49, 1305–1319, <https://doi.org/10.1007/s00382-016-3381-3>, 2017.

Li, J. Tree ring width based winter temperature reconstruction and its variation over the past nearly three hundred years in southwestern Sichuan Province, Master thesis, Henan University, 2021.

3. L38: It seems the manuscript uses “climatology” here and throughout the paper to mean “average weather” or “long-term averages.” It would help to specify the meaning of the term as it used here.

Response: Thanks for the comment. We use “climatology” to express long-term average in this manuscript. We have clarified it in the revised manuscript.

4. L40-50: The paper should expand and clarify this discussion of the strengths and limitations of conventional index methods and the possibilities of the Bayesian method, so that readers can better understand what is at stake. Some specific points to note:
 - The reconstruction skill for the conventional index methods is usually best for decadal scale variability (as discussed in Rudolf Brázdil et al., “Historical Climatology in Europe The State of the Art,” *Climatic Change* 70 (2005): 363–430 and Christian Pfister et al., “Analysis and Interpretation: Temperature and Precipitation Indices,” in *The Palgrave Handbook of Climate History*, ed. Sam White et al. (Palgrave Macmillan UK, 2018).)
 - Conventional index methods do not combine historical climatology and paleoclimatology. They use only archives of society.
 - Conventional index methods must leave gaps in a reconstruction where there is no suitable documentary evidence.
 - Reconstructions based on conventional index methods make one best estimate of the value of the target variable rather than a probability distribution. The historical sciences are starting to recognize that these probability distributions are important for understanding (environmental) history. (See e.g., Myles Lavan, “Epistemic Uncertainty, Subjective Probability, and Ancient History,” *Journal of Interdisciplinary History* 50 (2019): 91–111.)
 - Other methods may also incorporate both documentary information and paleoclimate proxies. The paper should also compare this Bayesian methods to past methods for integrating historical climatology and paleoclimatology together into temperature reconstructions (e.g., Jürg Luterbacher et al., “European Seasonal and Annual Temperature Variability, Trends, and Extremes Since 1500,” *Science* 303 (2004): 1499–

503). This may require a little more explanation of how the Bayesian method gets more information out of the sources than a single index value.

Response: Thank you for your constructive comments and helpful suggestions. We agree with your points regarding the limitations of conventional index approach, including their reconstruction skill, dependence on documentary sources, and the presence of temporal gaps. We also agree that, compared with conventional index approach, the Bayesian approach has the advantage of estimating probability distributions of different climate conditions rather than providing a single best estimate. In addition, it can incorporate “no record” information to update paleoclimatology. The limitation and advantages were both mentioned in the Introduction and Discussion Section, and your Comment 27 provided an insightful framework to discuss the comparison between different approaches. In the Introduction, we will more explicitly outline the limitations of different approaches in order to better emphasize the value and uniqueness of the Bayesian framework for readers. Add we will expand the discussion to systematically compare the Bayesian approach with other approaches, highlighting their respective advantages and limitations. This comparison will also be summarized in a table for clarity in the Discussion Section.

5. L55-59: I would recommend rewriting these lines for clarity and precision. First, it creates a lot of confusion to include the CWI as part of the “data” rather than the reconstruction “method.” The method in this study was to start by creating a winter temperature index, then to analyze those index results, and finally to use the index together with an ensemble of simulations to produce a reconstruction (i.e., posterior probability distribution). By treating the CWI as “data” it seems to be something that was “given” (the literal meaning of “data”) rather than something created through analysis of records. In fact, even the term “documentary data” may be misleading here. In this Bayesian approach, you are using the presence or absence of different documentary records as a means to update your beliefs about past weather and climate. Therefore, I would just say “documentary records.” Moreover, I wouldn’t say that the records “contain information about winter climate or meteorological events.” To maintain the strict logic of the Bayesian method, you should say that “winter climate and meteorological events shaped the contents of these records.” Causation always goes from the weather to the documents. The phrase “expressed through cold winter indices” is also confusing, since it suggests that the documents themselves included these indices.

I think I understand what you mean by the sentence “Paleo-simulations provided a climate background where all the information into which all information was assimilated.” However, the sentence literally makes no sense and it needs to be rewritten. The meaning of the following sentence (“The instrumental data is used in temperature-precipitation relationship analysis during the past decades and aided in determining

likelihood”) is also unclear. Just write out clearly what you mean here, even if it takes two or more sentences.

In section 2.1, I would again encourage the use of “documentary records” rather than “data.”

Response: Thank you for the comment and helpful suggestions. This paragraph is intended to provide a brief overview of how each type of data is used in the study. We agree that it would be better to mention CWI in the Methods section rather than the Data section. We will simply introduce documentary records without including the cold winter index and follow the logic of the Bayesian method. Regarding the terminology, based on a comprehensive review of the historical climatology literature, the term “documentary records” is more widely used and clearer than “documentary data”. We will replace this term throughout the revised manuscript. And we will revise the description of historical records to better align with the logic of the Bayesian approach and to clarify their dual role in both index generation and likelihood estimation. For the sentence about paleo-simulations, and we will revise this sentence to explain the role of Paleo-simulations as physically constrained priors. The introduction of the instrumental data will also be revised to clarify its role.

6. L68: The phrase “supplementation and validation” should presumably be “to supplement or to validate.” I assume the authors mean that some records were studied to provide additional information for the reconstruction (i.e., to supplement the reconstruction) while others were left out of the reconstruction in order to test the final result (i.e., to validate the reconstruction). This should be clear in the text.

Response: Thank you for the comment. We apologize for the lack of clarity in the manuscript. In Line 65, the “three compilations” refer to three compilations of documentary records covering the entirety of China, and thus also the full study area: *A Compendium of Chinese Meteorological Records of the Last 3000 Years*, *Integrated Natural Disaster Information System of the Qing Dynasty*, and *The Encyclopedia of Meteorological Disasters in China*. Historians and climatologists extracted and compiled weather- and climate-related records from original historical sources to these compilations. For the reconstruction period considered here, the first two compilations only cover the era before 1911, while the third extends beyond 1949. To avoid potential records between 1911 and 1949 which were not compiled to the third compilation, we therefore consulted original local gazetteers for this period. Two other compilations which only covers one province in study area are also used but provides less records. Above process constitutes what we refer to as “supplementation and validation” and it not clear enough for readers. We will revise this paragraph to improve clarity, and we will add a table to describe the type, time period, spatial coverage of all the sources in the supplement materials.

7. L105: I think “documentary information” or “documentary evidence” would be better than “documentary proxies.” Also, does ModE-RA use the same or different documentary records than those describes in section 2.1?

Response: Thanks for the comments and suggestions. For this sentence, we will revise the term “natural and documentary proxies” to “natural proxies and documentary evidence”. For ModE-RA, on the one hand, the project mainly collected the climatic indices from REACHES database generated by Wang et al (2018). The REACHES database was digitized and coded based on the *A Compendium of Chinese Meteorological Records of the Last 3000 Years*, which is one of our main sources but not the only one. On the other hand, the REACHES database was not fully assimilated into the final output in ModE-RA.

Reference:

Wang, P., Lin, K., Liao, Y., Lin, Y., Hsu, C., Hsu, S., Wan, C., Lee, S., Fan, I., Tan, P., Ting, T: Construction of the REACHES climate database based on historical documents of China. *Sci Data* 5, 180288, <https://doi.org/10.1038/sdata.2018.288>, 2018.

8. L106: Valler et al. 2024 does not actually use the term “time-independent.” Perhaps another term should be used here, or the term should be explained. (The explanation comes in lines 201–202, but that is too late.)

Response: Thanks for the comments. Valler et al. (2024) used the term “random stationary priors” to introduce ModE-RAclim, and we will revise the term and add explanation in the manuscript to make it consistent.

Reference

Valler, V., Franke, J., Brugnara, Y., Samakinwa, E., Hand, R., Lundstad, E., Burgdorf, A.-M., Lipfert, L., Friedman, A. R., & Brönnimann, S: ModE-RA: A global monthly paleo-reanalysis of the modern era 1421 to 2008. *Scientific Data*, 11(1), 36. <https://doi.org/10.1038/s41597-023-02733-8>, 2024.

9. L188: The phrase “were averagely deviated from climatology with around -0.5 standard deviation” sounds awkward. I bring this up as just one example of writing that the authors should correct and clarify before publication.

Response: Thank you for the comments. We apologize for the unclear wording in some sentences. We will revise these sentences and thoroughly review the manuscript to improve clarity and correct remaining issues with phrasing.

10. In Table I: The description for CWI 3 is “no record; the weather is as normal.” However, the absence of a record is only a possible indicator that the weather was normal. In theory, there could be no record even for very cold or mild winters. Therefore, I would recommend a revising the description. Perhaps: “normal weather; no exceptional events for observers to record”.

Response: Thank you for the comment and helpful suggestions. For the CWI 3, “no record” and “the weather is as normal” represent two different conditions in the original sources. We will modify Table 1 to clarify it and add an example phrase corresponding to “the weather is as normal” in the revised manuscript.

11. L204: The manuscript refers to both the CWI (in section 4.1.2) and the final posterior probability distributions of winter temperature (in section 4.2) as “the reconstruction” or “reconstructions.” Moreover, it uses the acronym CWI for the cold winter indices and then uses “CWI” again in the acronyms for each of the posterior probability distributions. This can be very confusing for the reader. I recommend that the manuscript stop referring to the CWI as a “reconstruction.” Moreover, the authors should find less confusing acronyms than CWI-ModE-Sim, CWI-LME, and CWI-ModEClim—for example, by switching “CWI” to “Post” (for posterior) (i.e., Post-ModE-Sim, etc.)

Response: Thank you for pointing this out. In the revised manuscript, we will no longer use “reconstructions” to refer to CWI. And we will adopt more precise acronyms to distinguish between the different reconstructions and reduce potential ambiguity. We will also include a schematic workflow to clearly illustrate the reconstruction process and its various outputs.

12. L206: “Likelihood assignment requires estimating the probability of observations being true given different indices.” Although this is a correct description, I recommend that the authors explain it more clearly. Because the likelihood estimate reverses the usual approach in historical research (i.e., estimating past conditions based on present evidence) it may be counter-intuitive, and some readers will misunderstand it.

Response: Thank you for the comment. We agree that this point may not be sufficiently clear. In the revised manuscript, we will expand the explanation of the likelihood estimation and better highlight how its logic differs from more conventional approaches to reduce potential misunderstanding.

13. L214: From this point on, the manuscript uses the phrase “likelihood of each winter.” This is a very confusing expression to a native speaker, as it would literally mean the chances of there being a winter in a given year. The authors should always spell out what they mean—i.e., the likelihood of getting the evidence for each state of winter temperature.

Response: Thank you for pointing this out and for the helpful suggestion. In the revised manuscript, we will not use ambiguous expressions of “likelihood of each winter” and instead clarify the meaning of the likelihood to reduce potential misunderstanding.

14. L229–231: This way this is written is confusing. Moreover, wouldn’t it make more sense just to provide a posterior probability distribution (similar to Camenisch et al. 2022)?

Response: Thank you for the comment. We will provide the prior, likelihood, and posterior probability distribution in the final Supplementary Material. We presented the best estimates and associated uncertainties of the temperature values based on the posterior probability distribution to facilitate comparison with other reconstructions in Section 4.2.2.

15. L235: Rather than immediately beginning with sub-sub-section 4.1.1, I recommend that the authors first briefly explain how section 4 will be organized (under the main heading for section 4) and then how the sub-section 4.1 will be organized (under the subheading for section 4.1). Section 3 has already gone through the entire method up to the steps for finding the posterior probability distribution for temperature. Therefore, the manuscript must explain to the reader that section 4.1 is going back to the first step in the method (the production and analysis of the CWI). Otherwise, it looks like the manuscript is heading directly into the final temperature reconstruction. This potential confusion that is exacerbated by the manuscript's confusing use of the term "reconstruction" for the CWI, as well as the letters "CWI" in the acronyms for the final reconstructions (i.e., the posteriors).

Response: Thank you for the comment and helpful suggestions. We recognize that parts of the methodological description are not clear enough, particularly given the complexity of the Bayesian framework. In the revised manuscript, we will add a brief introduction in Section 4 and in each subsection to clarify their organization. We will also revisit the relevant methodological steps when presenting the results, where appropriate, to better guide the reader.

16. L290–300: The authors should keep in mind that "accuracy" in conventional index methods simply means the probability that available descriptions are true representations of the weather for the season they describe. For likelihood estimates in the Bayesian method, however, we want to know both the sufficiency and necessity of each state of the weather to leave corresponding descriptions in our records. Some states of the weather (such as very cold winters) might be highly sufficient to produce corresponding descriptions but not strictly necessary (because some normal winters are described as cold). Conversely, some states of the weather (such as average winters) might be highly necessary to produce corresponding descriptions but not sufficient (because most observers didn't bother to record average conditions). Therefore, conventional terms such as "accuracy" and "reliability" may not be appropriate. It may be more appropriate to explain the value of records in terms of their power to update our knowledge about the past using a Bayesian method.

Response: Thank you for this valuable comment. We agree with you that, within a Bayesian framework, the relationship between weather/climate states and documentary evidence should be interpreted in terms of the sufficiency and necessity of different states of weather/climate to

generate corresponding descriptions, which we refer to as the dependence of observations on the climate state in the manuscript. In our likelihood estimation, we firstly used records from 1950–1999 as a reference to help characterize the relationship between climate states and documentary evidence. Second, we consider the uncertainty among different documentary sources, where records describing similar events may differ in reliability due to differences in origin; this uncertainty is evaluated following the framework of Ge & Zhang (1990). We acknowledge that terms such as “accuracy” and “reliability” may not fully align with the Bayesian interpretation. Therefore, in the revised manuscript, we will revise the explanation to better align with a Bayesian perspective, emphasizing the role of documentary evidence in updating prior knowledge about past climate states.

Reference

Ge Q. & Zhang, P.: The Evaluation of Climatic Information in the Historical Literatures. *ACTA GEOGRAPHICA SINICA*, 45(1). <https://doi.org/10.11821/xb199001003>, 1990.

17. In line 315, “the index method” should be changed to “all index methods” since the problem applies as much or more to conventional index methods as the Bayesian method.

Response: Thank you for the comment. We will revise it.

18. The image (a) in figure 4 is acceptable, but it might be redrawn to display more information and to present it more clearly. A 10-year moving average might be more informative than the decadal averages. Bars for annual CWI might be easier to read than the squiggly black line. Shading or color might possibly be used to express reliability (as described in 4.2.1).

Response: Thank you for the helpful suggestions. We will redraw the Figure 4a to make it more informative.

19. L337: Why does the heading for subsection 4.2 say that the reconstruction begins in 1750, when it begins in 1700?

Response: Thank you for pointing this out, and we apologize for this typographical error. The reconstruction indeed begins in 1700, and we have corrected this heading in the revised manuscript.

20. L337: Again, I would recommend that the authors include text under the sub-heading rather than beginning immediately under the sub-sub-heading 4.2.1. I recommend a couple sentences reminding readers of the meaning of “likelihoods” and their role in producing a final posterior probability distribution for temperature. Likelihoods were not explained clearly in 3.3.2, and therefore many readers will not understand the method here in section 4.2.1.

Response: Thank you for your comment. In Section 3.3.2, we briefly explained the meaning of “likelihood” and focused more on describing how it is estimated. We recognize the limitations of this expression. In the revised manuscript, we will add a clearer explanation of likelihood and its role in the whole Bayesian framework. In addition, we will add a short introductory paragraph under the Section 4.2 heading to remind readers of the meaning of “likelihood” and its role.

21. L379: For clarity, I recommend that the authors write out terms clearly. Here, for example, “posterior probability” should be “posterior probability distribution for each year’s winter temperature.” This approach will be new to most readers, and the manuscript should take pains to avoid vagueness or confusion.

Response: Thank you for the suggestions. We agree that many of the concepts and terms in the manuscript may be confusing to readers unfamiliar with this approach. We will carefully review this manuscript and will replace terms or concepts with more complete and clear ones.

22. L380–381: “Using the posterior, we reconstructed winter temperatures since 1700 and estimated their uncertainties”: This statement is confusing, because the posterior probability distribution for each year’s winter temperature is already a winter temperature reconstruction with uncertainties. It may be more accurate to state that the authors “used the posterior to create figure 6” or “used the posterior to visualize the best estimate and range of uncertainty for winter temperatures since 1700, as shown in figure 6.”

Response: Thank you for your comment. Within the Bayesian framework, we first reconstructed the posterior probability distribution of the cold winter index for each winter. Based on this posterior probability distribution, we then calculated the best estimation of temperature values and quantified their associated uncertainties. These results were further visualized and analyzed in Section 4.2.2. In the revised manuscript, we will rewrite the relevant sentences to make the methodological process and its presentation clearer.

23. L385–386: “In the 18th century, our reconstruction results difference about the warming in the 1760s.” This sentence makes no sense. Please re-write.

Response: Thank you for pointing this out. We apologize for the unclear expression. The intended meaning of this sentence is that our reconstruction differs from Yang’s reconstruction in the warming trend during the 1760s. We will revise the sentence accordingly in the manuscript.

Reference

Yang, Y.: The Preliminary Study in Reconstructing the Mean Winter Temperature in Kunming during 1721-1900 A. D. *Journal of Chinese Historical Geography*, 1, 17–31, 2007.

24. Figure 6: In addition to the reconstruction of each winter’s temperature, it might be interesting to compare the distribution of CWI values in figure 4(b) with the reconstructed

(posterior) frequency of winters in each range. This would further illustrate how the Bayesian approach can provide a more realistic depiction of interannual temperature variability than indices alone, particularly when the documentary information is less complete or less certain.

Response: Thank you for this interesting suggestion. We will further explore this comparison and include it in the revised manuscript if it provides any insightful results.

25. L406: The subsection title “4.3 Source of information: observations and simulations” is vague. The subsection title and the first sentences of the subsection should explain more clearly why the authors are comparing CWI-ModE-Clim to ModE-Sim.

Response: Thank you for the comment. We will revise the subsection title of 4.3 to make it more clearly reflect the content of this section. And we will add a brief paragraph in the beginning of this subsection to explain the meaning of the comparison between CWI-ModE-Clim and ModE-Sim.

26. L417-432: While this comparison between ERA5 and ModE-Sim is interesting, it's not clear from the text what this has to do with the Bayesian reconstruction.

Response: Thank you for the comment. We agree that the connection between this comparison and the Bayesian reconstruction was not sufficiently clear in the text. By comparing ERA5 and ModE-Sim, we assess whether the large-scale atmospheric patterns associated with cold and warm winters in the study region are consistently represented in the paleo-simulations which served as prior in this study. The results indicate that these patterns in paleo-simulations are broadly consistent with those in the ERA5 dataset, which supports its use in the Bayesian framework. We have added more explanation in this section to make this connection explicit.

27. L446–462: When discussing the “advantages” and “limitations” of the Bayesian approach, it is important to specify what you are comparing the Bayesian approach to. The Bayesian approach has advantages and disadvantages compared to conventional historical climatology as well as advantages and disadvantages compared to paleoclimate re-analyses. Perhaps it would be helpful to present a small table comparing key features of (1) Bayesian historical climatology, (2) conventional historical climatology, and (3) paleoclimate re-analyses. Features could include, e.g., ease of use, reconstruction skill, data requirements, objectivity/subjectivity, and accurate representation of uncertainties. In any case, this section should be revised to make it clear which of the “challenges” of the Bayesian approach are already present in conventional historical climatology and which of the “advantages” are unique to the Bayesian approach.

Response: Thank you for the comment and helpful suggestion. The suggested framework for comparing the Bayesian approach with conventional approaches is very insightful. In the

revised manuscript, we will add a summary table in the Discussion section to present this comparison more clearly. Specifically, we will compare the Bayesian approach with conventional approaches in terms of ease of implementation, the extent to which kind of historical records are utilized, and the representation of uncertainty, among other aspects. This comparison will help clarify which advantages and limitations are specific to the Bayesian approach and which are shared more broadly with conventional methods.