

Interactive comment on “Palaeo sea-level and ice-sheet databases: problems, strategies and perspectives” by A. Düsterhus et al.

Anonymous Referee #1

Received and published: 1 August 2015

General Comments:

This manuscript begins to discuss desirable qualities of a sea-level and ice-sheet database. The database itself would have great value, and this paper could be a first step towards database development. However, the paper needs considerable work in order to be truly useful. In general, it is too vague and does not provide enough discipline-specific information (which the PALSEA2 community certainly has access to). Throughout much of the paper, it is unclear what are general characteristics of a good database versus characteristics that are necessary for the paleo community. The former is not interesting here— that information is already established and presenting it is not the goal of the paper. Focus on discipline-specific material. In the comment sections below, I discuss various areas that could be made more specific.

C1219

The manuscript reads somewhat like a pre-emptive justification for a future grant proposal. It should have more value than that, and will indeed be more valuable to future proposals if it is more explicit in its discussion of problems and strategies.

The manuscript is also disorganized, with little apparent adherence to the section headings. (Specific areas that require reorganization are mentioned in the comments section below.) It sometimes feels like several authors wrote bits and pieces, then cobbled the pieces together. Early on, the authors state that a database must recognize three distinct phases. I suggest reorganizing the paper into sections on these phases— measurement, documentation, interpretation— as a guide to discuss more community-relevant details. Problems with each phase could be fleshed out in separate sections. This is just a suggestion. But, regardless, the paper needs reorganization in some manner or another.

Note: One way to provide clear, detailed strategies for database development would be through use of additional tables and/or figures, especially flowcharts. I suggest including several.

Second note: I'm not sure which Climate of the Past article category this falls into— Technical Notes? If so, it seems even more necessary to incorporate a more detailed treatment of database creation strategies.

Other Comments/Questions:

The authors briefly discuss the existence of other sea-level and ice-sheet databases, but might also benefit from looking at other geologic subfields. There are plenty of tectonic (e.g., <http://www.ig.utexas.edu/research/projects/plates/>), paleontological (e.g., <http://fossilworks.org/?page=paleodb>), petrological (<http://www.earthchem.org/petdb>), etc. databases that deal with complex problems akin to those discussed here. What has or hasn't been successful with these types of large, geologic databases?

What time scale are the authors discussing? A Quaternary database will have slightly

C1220

different needs than a Miocene-to-modern database, etc.

I think the paleoclimate community would welcome a good, comprehensive database such as that proposed by the authors. Therefore, we need little convincing. I would spend less space on the motivation behind developing a database, and more space—probably much more space—on exactly how to structure the database.

This manuscript is chockablock with vague imperatives (“essential that”, “important to” . . .). While well-meaning and generally true, their ill-defined nature makes them not very helpful. The (potential) value in this manuscript is that it can provide a starting point for database development. Please be as specific and directed as you can, and avoid the rather too obvious or too vague statements such as “it is important to consider the needs of the end-users”, “Global quality-controlled databases are necessary for answering the challenging questions about the Earth system,” and “can be handled in more transparent ways.”

If one of the primary goals of the database would be to provide input to sea-level and ice-sheet models, the authors should spend some space discussing exactly how to do that. What sort of database infrastructure would be especially useful for modellers, especially non-specialist modellers? Again, more DETAIL is required here. A list would be helpful.

Detailed Points:

(P = page, L= line)

P 2391, L 1: They are not essential, they are only helpful. If they were essential, then nothing that didn't use a database would be worth anything.

P 2391, L 3: Exchange “related” with “specific”

P 2391, L 3-4: Reword to use similar verbiage between phrases (e.g., “the composition and needs” and “finally using it”).

C1221

P 2391, L 4: “There are also issues” is too vague—do you mean discipline-specific issues?

P 2391, L 25: The manuscript only starts to touch on “strategies for developing a standardised database.” If this is the authors' goal, then do it!

P 2392, L 2-3: “Palaeo” here is unnecessary because you later say “point in time and space”. Also, watch the consistency of your use of “palaeo” versus “paleo”.

P 2392, L 3: Exchange “In particular” for “In addition”

P 2392, L 6: “a priori” requires italics.

P 2392, L 5-21: This discussion seems out of place. This is a fine treatment of some uncertainties dealt with in the paleo community, but does not fit in the introduction. Such a long discussion makes it seem like the uncertainties are really the most important part of database development. Is this true? If it is not, then put this discussion elsewhere. If it IS the most important part, then the paper needs to be structured around a treatment of uncertainties. . .

P 2392, L 21-23: Starting with “Finally”, this is a new topic. Is this introduction supposed to be a summary of the problems involved in a potential database for the paleo community? It is too disorganized for me to tell. One possible way to improve this would be to use the introduction to describe HOW the community does/could use databases. Alternatively, you could use the intro as a summary of problems, but you must then address more than just examples of uncertainties. Note that the ones listed in lines 5-21 are by no means exhaustive.

P 2392, L 25-29: This could be moved up earlier in the introduction.

P 2393, L 12: Remove “s” from end of “databases”

P 2393, L 15-16: Do you build on existing guidelines? It is unclear from the manuscript text. The Shennan and Balco papers are not referenced at all in the text (and the Balco

C1222

paper is not in the references).

P 2393, L 16-17: What are the key components? It is not clear. This would be an area that could use a table.

P 2393, L 19: Consider inserting “of geological records” after “databases”, or something similar.

P 2394, L 3-14: This seems out of place. Is this truly an aspect of the community structure? It seems more like suggested changes to funding structure, which might fit better toward the end of the paper. Alternatively, it could be incorporated into a section on paleo-community problems.

P 2394, L 15-17: Seems a bit bold. What other data are you comparing this to? The paleolimnologists, etc. might argue back.

P 2394, L 19-23: This is more discussion of uncertainties and should be combined with P 2392, L 5-10. Also, seems like this is a discussion of interpretation rather than standardisation.

P 2394, L 24: Are the authors suggesting that in order to have a standardised database, researchers should have standardised field measurements?

P 2395, L 3-4: Seems like this should be right up at the front of the manuscript.

P 2395, L 5: “Therefore” is out of place.

P 2395, L 5-8: Needs refs.

P 2395, L 8-10: Overlaps with P 2392, L 20. Consolidate the discussion of problems and uncertainties.

P 2395, L 11-18: Is this data reporting or database infrastructure? It seems to fit better in the current Section 4.

P 2396, L 2-8: Good! Some clarity and specificity. Do more of this.

C1223

P 2396, L 10: Insert “paleoclimate” or “geologic” or something before “databases”.

P 2396, L 13-19: Do you suggest using one-to-one or relational databases? From your later discussion, it seems like one-to-one would not be sufficient. Go beyond the statement that there are different kinds of databases and propose a STRATEGY, as promised in the abstract, of how the PALSEA2 community thinks the database should be structured. At minimum, perhaps a short table of pros and cons of a one-to-one (or spreadsheet) versus a more complex database could be useful.

P 2396, L 16-17: This sentence seems randomly placed. Why is it here? Is there something special about LGM databases?

P 2396, L 24: Need a transition between these sentences.

P 2397, L 1-12: This is an introductory paragraph and should be placed in the beginning.

P 2397, L 25-26: What are your suggested “consistent data types per column”? This is another example of a place to suggest strategies based on the PALSEA2 collective knowledge.

P 2398, L 6-7: This sentence is out of place and too vague to be meaningful.

P 2398, L 15-17: Good point. What should be done to address this? Require citations of the database to include a “date last modified”? Preserve old versions of the database for reference?

P 2398, L 18-21: Reword—this is unclear.

P 2398, L 9 to P 2399 L 9: How unique are these problems to the paleo community? Not very...? Have any other subfields addressed these problems? I agree that the problems raised are real and important, but it seems like this is a science-wide problem. This manuscript may not, therefore, be the place for a long-ish discussion of the problem. What IS important to address here is how the paleo community can develop

C1224

a dataset that overcomes (or starts to overcome) the problem of citation. It is not helpful to list general problems without providing the PALSEA2 community's insight into potential solutions.

P 2399, L 11-12: I am hoping that if the authors are putting in the effort to write a paper about a potential database, that it would indeed end up on the internet. That itself seems too obvious to include—however, if the authors wish to include such a statement, I would also make mention of access issues. I suspect that many young researchers, or researchers from outside of the country of origin (especially where language differences exist) do not know that of the existence of databases that are only acquired by contacting an author. I can also imagine instances when such researchers would hesitate to contact an intimidating bigwig scientist. Therefore, and especially when considered through the eyes of government funding agencies, I believe online access to be imperative.

P 2399, L 20: How do you suggest using this “information as a basis for future databases”? Be specific.

P 2399, L 22-23: Outline this “database of databases” for me. Which databases? Exclusively the ones in the supplementary table? Which data would you use? I'm looking for more detail as to your suggested strategy.

P 2399, L 24-25: This is more fitting for the very beginning or the conclusion.

P 2399, L 25 to P 2400, L 6: This seems to fit in more with the infrastructure section.

P 2400, L 12-14: I cannot support this statement. What are your new advances? I truly do not know. This is a strong indication of the disorganization and vagueness of the manuscript. Please make a list of the advances. Also, you need references to show that you have made “substantial progress”. Progress over what? It is difficult to determine what the manuscript is about, never mind about determining what progress the manuscript makes.

C1225

P 2400, L 22: What are the “goals defined herein”? Please list these.

P 2400, L 26 to P 2401, L 2: OK, what is the PALSEA2 community doing, then? Consider expanding on this statement to provide more detail.

Table: I like the idea, but this needs to be expanded on. First, improve the consistency of detail between entries. Second, it would be useful to see what kind of data each database provides.

Interactive comment on Clim. Past Discuss., 11, 2389, 2015.

C1226