

General comments

This objective of this paper is to use simulated paleoclimates over the past glacial/interglacial cycle to drive biome simulations, testing those simulations with the pollen data, and then applying the biome simulations to a simple terrestrial carbon/ $\delta^{13}\text{C}$ model in order to explain ocean $\delta^{13}\text{C}$ variations over time. The paper used the results of simulations from two climate models, biomes simulated from those climate-model outputs, and the “observed” record of biome variations inferred from fossil-pollen data to check to check the simulated biomes, and finally, calculations of terrestrial carbon-storage variations that in turn govern those $\delta^{13}\text{C}$ variations.

There is considerable fuzziness in describing which model does what here, and the difference between simulations and reconstructions. The two climate models simulate climate, but not (as they are used here) elements of the carbon cycle (i.e. “interactive vegetation is not included” (in HadCM3), and “interactive vegetation was not used” (in FAMOUS)), so expressions like “The two climate models show good agreement in global and net primary productivity...” don’t make sense. Also, the output of BIOME4 is simulated vegetation, as opposed to reconstructed vegetation, which is the product of the pollen synthesis. The general experimental design applied here of using simulated climate to drive biome simulations, testing those simulations with the pollen data, and then applying the biome simulations to a simple model $\delta^{13}\text{C}$ model emerges slowly in the paper (with the last step not really being discussed until 31 pages into the paper), and so the design might usefully be stated in a “here we use...” fashion in the abstract to get the reader off on the right track.

The data-model comparison (between simulated and “observed” biomes) is relatively lightweight, featuring only a few map comparisons with the relative sparse network of “long” records used here. I was expecting a comparison involving the full BIOME6000 0, 6 and 21 ka data set (which would not be hard to do). It is asserted that the hierarchy of models can successfully simulated the biomes inferred from the pollen data, but this “working hypothesis” as it’s called, is never tested, although it could and should be.

There are three potential sources of data-model mismatches (or indeed accidental matches), 1) the model, 2) the data and 3) the experimental design. For example, disagreements between simulated biomes and observed (“biomized”) biomes could be attributable to the hierarchy of models (AOGCMs and BIOME4), the biomization process itself, or to the experimental design (in the application of all of the models). It would be good to discuss those sources and the extent to which each could be influencing the results here. Overall, I don’t think the case has been made that the hierarchy of models works well enough to use the BIOME4 output for carbon-budget calculations. That the approach does work could be demonstrated using the BIOME6000 data, along with multi-model simulations of 6 and 21 ka; this would also help to evaluate the relative importance of the three sources of mismatch.

There is a relatively large author list, which includes some but not contributors of the original data, and which overlaps a lot with the authors of the individual papers in the Sanchez-Goñi and Harrison (2010) QSR special issue on millennial-scale climate variability and vegetation changes during the last glacial-interglacial cycle. It would be appropriate to provide an indication of author contributions.

The figures need some work. The key figure is Fig. 2, which shows simulated and observed biomes, but fuzzes up at the scale necessary to view the results for individual continents. Zooming way in on Fig. 3 suggests that the curves may be “spikier” than they should (i.e. in the data), because it looks like they were constructed with “bevel-joined” line ends (which extrapolate the data, creating the sharp spikes), instead of the more appropriate, but inelegantly named, “butt-joined” line ends.

Specific comments:

p. 1034/line 5: replace “Global ...” distributions” with “Simulated (BIOME4) biome distributions at the global scale” (or something like that).

1034/9: “modelled changes in vegetation” I think this should read “simulated changes in vegetation”—the modelling work got done as BIOME4 was developed; here the model is being applied to generate simulations.

1034/25: “Quasi-periodic” What’s quasi about the periodicity?

1035/3: “...for the last ~0.8 million years...” What’s special about that interval? Orbital variations have never not influenced climate (and the biosphere, after it developed). It might be better to review the particular variations of climate and its controls over the last glacial cycle than to describe the general ice-sheet, sea-level, CO₂, etc. (Quaternary 1010) relationships.

1035/9: “productivity and size of the terrestrial biosphere” “Size” could be interpreted a number of different ways, including areal extent, total biomass, etc.

1035/14: “... the terrestrial biosphere was significantly reduced as forests contracted.” Reduced in what sense? I think the area of the terrestrial biosphere varies rather little over time as ice-covered areas seem to be roughly compensated for by exposed shelves. Does this mean instead that forested areas were reduced in area?

1035/15: “21 kaBP” means “21,000 years ago before present”. Just “21 ka”.

1036/4: “The data can be viewed through the prism of a global, physically based model that allows the point-wise data to be joined together in a coherent way.” Does that simply mean “interpolation” (which you’re not doing here). Or are you describing how to comparing a sparse network of reconstructions with gridded simulations? In any case this sounds like text from a proposal as opposed to a description of what was done here.

1036/6: “There are continuous, multi-millennial palaeoenvironmental records... that have not been previously brought together is a global synthesis.” Given the author overlap between this paper and those in the Sanchez-Goñi and Harrison (2010) QSR special issue, this statement is a little surprising. Also, only one kind of palaeoenvironmental data is being synthesized here.

1036/14: “We present quantitative estimates of changes in the terrestrial biosphere reconstructed from two atmosphere-ocean general circulation model (AOGCM) simulations over the last glacial cycle.” No you don’t—the “quantitative estimates” come out of BIOME4.

1036/22: “We assess...” There’s a step missing here. How are biome simulations turned into $\delta^{13}\text{C}$ values? (Actually the biome-simulation step is missing too.)

1037/3: “Biomization assigns ... based on biological and climatological ranges.” To a reader unfamiliar with this process, that might sound like some kind of calibration with climate data is involved.

1037/16: “megabiome score data...” Why are there blank rows in the spreadsheets? For example, there are pollen data for the Carp. L. sample at 6.12m, but no (mega)biome scores. (Also, why are there two age models for this record?)

1038/4: “reconstructions” again

1038/8: “climate averages” “long-term monthly means”?

Sections 2.2.2 and 2.2.3: How was land-surface cover specified (or calculated) in the simulations?

1039/10: “biogeochemistry-biogeography model” Should that aspect of the model be mentioned earlier.

1039/20: “compare well with NGRIP...” Were similar comparisons done for HadCM3?

1039/24: “physically justified ice-sheet extents” Explain.

1040/20: “adjusting ... to compensate for ... biases” and (line 22) “Climate model anomalies ...” Is this two separate steps (bias-correction, and then the calculation of anomalies)? What was the base period for the anomalies?

1040/22: “temperature and precipitation” What about sunshine, and how was changing insolation handled?

1040/23: “Leemans and Cramer” This implies that BIOME4 was run over the 0.5-degree grid of this data set, but Fig. 3 shows simulated biomes on the grids of the AOGCMs. There’s a big step missing here.

1040/27: “model’s” Which one?

1041/1: “no special correction...” How were modern climate values created for the exposed shelves?

1041/4: “BIOME4 was forced with appropriate CO₂ ... (same as used to force the climate model)” Does this mean that Vostok CO₂ was used for the HadCM3-driven simulations and EPICA CO₂ for the FAMOUS-driven simulations?

1042/8: Southeastern? (also in line 15, San Felipe and Potato Lake would commonly be located in the Southwestern US).

1042/27: “Recent” as in “present day” or newer than Thompson and Anderson (2000)?

1042/29: “... those of the LGM also compare well.” With what?

Section 3.1: I’m not sure this section serves the paper very well. Each subsection starts with an overview of the location of the sites, but then rapidly becomes anecdotal, describing some aspects of the record for some sites, and different aspects for others. One overall impression I got is that the biomes don’t vary much over time, and another is that there are important differences between the (mega) biomizations here and what was produced in previous studies; neither impression increases confidence about the results. This is likely more a consequence of the way this section was written than of real issues in the data. As the authors indicate (p. 1041, lines 14-15) only the main results are being presented, but there is no overarching summary—the paper just moves on to the simulation results. The paper promises a new synthesis, but all it delivers is a few dots on Fig. 2, and some spreadsheets that list the affinity scores, but not the actual reconstructed biomes that the paper is based on.

I wonder if this section could be moved to supplemental information, where a more systematic discussion of the individual records could be done, and replaced in the main text with some kind of summary figure. Alternatively, the reader could simply be referred to the Harrison and Sanchez-Goñi summary article in the QSR issue, along with the individual regional articles in that issue. I’m going to skip commenting on the reset of this section.

1048/18: “where they disagree...” This paragraph starts out talking about the source codes of the climate models, and so it would be easy for the reader to surmise that the disagreement mentioned here is between the climate models and not between the BIOME4 simulations.

1040/21: “coupled to BIOME4” That’s not really happening here.

1049/6: “Because of its lower resolution...” This is certainly true at the resolution of the GCMs, but earlier the experimental design was described as including the “apply-the-anomalies” approach to the 0.5-degree Leemans and Cramer data set (and repeated on p. 1050/line 6), so presumably the modern “high-resolution” spatial climate variations are also present in the input data for BIOME4.

1049/13: “difference in temperature...” When? At present, or over the course of the climate simulations? (Same question for precipitation...)

1049/21: “warm bias” Again, when?

1049/21: “Millennial-scale cooling events...are not features of our model runs...” Does this mean that they were not simulated, or that the experimental design of the climate simulations did not include the appropriate forcing?

1049/28: Replace “modelled reconstructions” with “simulated biomes”.

1050/4: The caption for Fig. 2 should point out that it shows both simulated and reconstructed megabiomes. The caption should also explain what we’re seeing on the grids of the two models. One simulated biome at the grid point? The modal simulated biomes (on the 0.5-degree grid) within the area represented by each model grid cell?

1050/19: “additional warmth and sea level” “higher temperature and sea level”?

1051/5: “Differences between our pre-industrial megabiome reconstructions [read “simulated biomes”] only arise from the way the pre-industrial climate forcing [Leemans and Cramer, right?] has been interpolated onto the model grids.” How was the Leemans and Cramer data interpolated onto the model grids? More to the point, why was it necessary to do that? The anomalies of the pre-industrial simulated climate relative to themselves are zero, so there shouldn’t be any difference in simulated biomes, unless something that hasn’t been explained is going on.

1051/25: “... on the scale of the climate-model gridboxes.” This makes me think that the biomes are being simulated only for each model grid point, and not for each 0.5-degree grid point in the Leemans and Cramer data set via the “apply-the-anomalies” approach. If that’s the case, the poor agreement throughout between simulated and reconstructed biomes makes sense.

1052/9: “... this comparison gives reasonable support to our working hypothesis...” That hypothesis is testable, and indeed should be. It looks to me like there are as many sites with inferred biomes that differ from the simulated biomes as don’t. If you can’t convincingly show that biomes inferred from “modern” pollen data match those simulated by observed climate, then why should we believe the results for other times?

1052/13: “For both the mid-Holocene and LGM periods, the high-resolution biomizations of the BIOME6000 project (see Table 1) provide a better base...” The same is true for the present.

1052/19: “a greening”?

1052/24: “weak precipitation”?

1052/25: “FAMOUS shows a smaller reduction”?

1052/27: “regional biome reconstructions” Do you mean yours here or the BIOME6000 ones?

1052/28: “magnitude of the rainfall” The magnitude of the rainfall or of the rainfall anomaly?

1053/5: “wetter anomalies” Wetter than what? (And it would be better to talk about changes in precipitation as opposed to “wetness”.)

Sections 3.3.2 and 3.3.3: As was the case for the present day, it looks to me that (in the absence) of any quantitative measures, there are as many disagreements and agreements. It is asserted that the more abundant biome data from BIOME6000 shows that “... there is again good general agreement between the two different model reconstructions and the regional biomizations of the BIOME6000 project.” (p. 1053, line 27) but there’s no real evidence that such is the case.

1055/11: “The similar model-based reconstructions...” Similar to what? The LGM simulations? The biome reconstructions?

1055/23: “realistic two-dome pattern” Citation?

1055/26: “limited vegetation extent” “Vegetation” in this paragraph seems to be equated with tree cover here, but was just used above in the context of land not covered by ice.

1055/27: “wetter climate in HadCM3” Wetter than what?

1056/1: “cooler in FAMOUS” Cooler than what?

Sections 3.3.3-3.3.5: The main “take away” message I get from these sections is that there is almost no change in simulated or reconstructed biomes over this 40,000 year-long interval, with the ice/land mask accounting for most of any change in the simulations. Is that right?

1056/25: “similar affinity scores to the 64 ka..” “similar affinity scores to those at 64 ka”?

1056/26: “they are sparse” Sites in general, or those with similar affinity scores at 64 and 84 ka?

1057/1: Warmer than what?

1057/7: “poorly modelled Mediterranean storm-tracks...” What’s the basis for that assertion?

1057/9: “Although there are still differences....” I don’t understand “still”

1057/12: “larger areas of forest”?

1057/13: “a dry anomaly ... that reduces vegetation” Again I think you’re equating “vegetation” with “forest”.

1057/25: “regional climate feedbacks” Explain.

1057/28: “in line with ... each other” I’m not sure what this means. Simply “both models”?

1058/3: Both models increase the extent of their tropical forests...” Does this refer to the BIOME4 simulations? Throughout this section the climate simulations from the GCMs and the biome simulations from BIOME4 keep being conflated.

1058/11: “Quantitative estimates ... can thus be drawn...” Yes, but are they meaningful?

1058/16: “their overall effects” Overall effects of what? From proximity, “their” would refer to “areas and periods with significant regional differences” but that doesn’t make sense.

1058/18: Fig 3. There are three curves shown in each panel. I’m guessing “_S” means shelves and “_NS” means no shelves, but this isn’t explained

1058/23: “The changes in atmospheric CO₂ levels ... are common to all BIOME4 runs.” Two things: 1) CO₂ changes over time, so it makes no sense that the same levels were used for all runs. 2) CO₂ levels presumably don’t vary within a single simulation.

1059/3: “FAMOUS also neglects the additional area of land ...” What’s the argument here?

1059/5: “global total areas of biomes”?

1059/14: “several sites (Fig. 4)” Curves for only one site area plotted in Fig. 4.

1059/16: “~ 70 to 75 PgCyr⁻¹” Which is which? (Later you discuss the NPP values simulated by the two different simulated climates.)

1059/23: “... BIOME4 is driven solely by an observational climate dataset...” (for the PI), so on line 28, the “lower resolution topography” being referred to is that in the 0.5-degree data, right? I find it hard to believe that there is enough smoothing in those data (relative to elevation in the real world) to account for all of the positive NPP bias.

1060/6: “In the LGM simulations....”

1060/20: “Further analysis with HadCM3 suggests...” What kind of analysis?

1060/28: “Some differences in the timing of some events... are apparent...” What events?

1061/7: “lower NPP” Than what?

1061/22: Prentice et al. (1993). Not in references.

1062/13: Wang et al. (2011). Not in references.

1062/11 – 1063/2: This discussion is really methods, not results. Are the turnover times for different biomes tabulated anywhere? Where does the exponential decay multiplier come from?

1063/3: “The differences in modern NPP by biome between HadCM3 and FAMOUS (related resolution differences...” Please explain. Is “modern” different from “PI”? If so, there’s a whole set of simulations that haven’t been described anywhere (see also comments about p. 1051). If not, why should there be differences?

1063/14: “greater retention” Retained from what? (Sounds like from present...)

1063/23: “greater level of periodicity” I think you’re confusing the amplitude with the presence or absence of variations at the ~ 23 kyr time scale.

1063/26: “For the biome scores ... (Fig. 3).” Figure 3 shows simulated biome areas.

1063/28: “The largest impact...” On what? The areas? The periodicity of the variations?

1064/8: “... because other forest types are not compensating periodicities in grassland variation...” No idea what this means.

1064/15: I wonder at this point how much of the variation in Fig. 5 is related to the differences in the simulated climates and how much to the turnover times.

1064/17: This section has a lot of methods in it, and is rather late in the paper.

1065/10: “by the model output $\delta^{13}\text{C}$ for each grid cell” Where do those values come from?

1065/14: “did not estimate $\delta^{13}\text{C}$ values” “did not vary (atmospheric) $\delta^{13}\text{C}$ values”?

1065/16: “the calculated $\delta^{13}\text{C}$ ocean changes would not change” “would not vary”?

1065/20: “total ocean $\delta^{13}\text{C}$ was calculated for the last 120 kyr (Fig. 6b). Fig. 6b look like it shows anomalies from present day.

1066/ 18: “FAMOUS variation is nearly twice the magnitude” Twice the amplitude?

1066/21: “deep Pacific $\delta^{13}\text{C}$ records” Where are those shown?

1067/21: “Estimates of global carbon storage reduction are significantly greater if continental shelf exposure is not included...” But the shelves were exposed, so I’m not sure why this is even worth talking about.

1068/4: “regional climate biases” Biases weren’t ever assessed. The simulated climates differ from one another, but they were never compared with climate reconstructions.