I was looking forward to reviewing this paper because (a) it is exactly the sort of paper that we hoped the existence of GMD would enable and encourage when journal first started, (b) I have been involved in a similar paper describing our own model (Valdes et al, 2017), and (c) it has a section on configuring the model for deep-time simulations, which is of great interest scientifically to me.

I think that this paper will be very useful, providing a definitive benchmark for this model. I hope that the authors start a Special Issue on this model, and submit bug-fixes/developments as they occur, in follow-up papers.

I have some questions/comments (none of which are major), some suggestions for additional work (which I would view as nice but optional), and some minor corrections.

Dan Lunt

Questions/Comments:

- Introduction: as well as paleoclimate, there will be interest from the forcings/feedbacks community associated with longrunMIP (e.g. Rugenstein et al, 2019). Would be good to mention this. In general, I think you underplay the interest of this model to the non-paleo community.
- Abstract needs to include the resolution of the model (horizontal and vertical, atmosphere and ocean).
- Line 20: Fully complex (e.g. CMIP6) models don't only carry out snapshots, often transient runs (e.g. RCPs!). More generally, remember the audience for this paper is not just paleo scientists - you should assume an audience that is not familiar with paleoclimates, e.g. avoid terms such as "Quaternary", "last interglacial" without explaining when they were.
- Line 64: I would remove the statistical approach from the transient section, and instead add another two types of approach here, which are:
  (a) "multiple snapshots". Within this you could include e.g. Singarayer and Valdes (2010), and Marzocchi et al (2015).
  (b) statistical approaches based on multiple simulations, e.g. Lord et al (which you already cite), and Simon et al (2017), and Erb et al (2015).
- Line 96: It is not always clear from the description which components are identical to IPSL-CM5A, and which are "updated versions of IPSL-CM5A components"
- Line 116 indicates that "The differences mostly include bug corrections, various improvements in the physics/dynamics interface, and energy conservation. Changes also concern the Emanuel convection scheme, with bug corrections and changes in the upper bound of the deep convection loop.". it would be good to maybe say a little bit more about these, and explain why these changes were made.
- Figure 1 looks pretty, but would include more information if a standard cylindrical projection were used? Or maybe if views towards the Pacific were also added.
- Section 2.1: I like the way that the actual tagged releases of each component are given for the new model and for the old model. This is very good practice. Is there also a version number for the overall model, for both new and old versions? And is there a unique update number for NEMO and OASIS as well (the later section on Code Availability indicates that there is)?
- Line 229: Not sure what "adjusted" means in "provided an adjusted global surface air temperature of 11.3°C"
- Line 230: "Such a cold bias had been depicted in the previous IPSL-CM5A (Dufresne et al., 2013)". What was the magnitude of that bias?
- Line 231: "We targeted to increase global-mean surface temperature by 2.2°C to reach 13.5°C in pre-industrial conditions with IPSL-CM5A2, expecting this value to translate into 15.5°C in simulations with present-day conditions.". This needs a reference to some observational dataset, either to validate the 13.5 preindustrial choice, and/or the 15.5 present-day choice. Also mention the uncertainty in these observations.

- Line 236: “Previous results obtained with LMDZ and IPSL-CM5A show that changing TOA balance by 1 W.m−2 results in a change in temperature of 1 K.” I guess you mean “initial TOA inbalance”? Is this the TOA in a PI run, or historical, or….?
- Line 237: “It is also the typical value of the sensitivity of global temperature to greenhouse gas concentration” is rather opaque….do you mean that the overall net feedback parameter, lambda, in this model is ~1.0 W/m2 / K ?
- Line 249: “when two of the bugs mentioned above”…there are several bugs mentioned above…which ones are you referring to?
- Line 257: “respectively 0.029 W.m-2 and 0.023 W.m−2”. Is the difference between these two primarily due to interannual variability, or gradual accumulation of energy in the ocean, or bugs in energy conservations, or something else?
- Line 258: What are the causes of the non-conservation of energy?
- In very long simulations, small non-conservation of energy is not a problem, because the energy system is “open” to the outside – i.e. a small non-conservation may have very similar results to slightly increasing the solar constant – the model just finds a new equilibrium. However, the hydrological system is “closed”, and so non-conservation can result in substantial drifts in e.g. global mean salinity over very long time periods. How well does the model conserve water, and are any correction factors included to maintain global mean water/salinity content? OK, I see this is addressed later (lines 272-280). Did you consider a salinity correction term? See e.g. the HadCM3 runs in DeepMIP – Lunt et al (submitted to Climate of the Past).
- Line 268: “Such a difference is also small enough compared to the regional biases in sea surface temperature.”. How do you define “small enough”?! Would 0.5 degrees by small enough? What about 0.8, or 1.0?! Better to compare with some other models, and just say, e.g. “this is closer to observations than X% of CMIP5 models”, or (and) link it to the uncertainty in the observations – 0.3 degrees is probably within error of the observations?
- Supp Info could be used for some more plots, e.g. the old model version of Figure 4(d,e), the old model version of Figure 5(c). 6(a,d) etc etc. I expect you have already made these plots!
- Label all figures, not only in the caption but above the figure panels as well. Some plots are model minus observations, some model minus model, some just model. Clearere labelling would help (I know I could get this info from the caption but this is much slower than labelling the panels!)
- Fugure 15a – presumably this is the new model (not stated in the caption). Would be good to see this also for the old model.
- Line 595: would be good to show evolution of TOA flux, and show a Gregory plot.
- Section 5.4. This section implies that the model worked “first-time”, “out-of-the-box” for the new Cretaceous paleogeography. I would be surprised (and impressed!) if this were the case. If not, then it would be good to explain what needed to be done to allow the model to run, e.g. smoothing topography/bathymetry etc.

Suggestions

- Line 282: Presumably there is quite a “kick” to the model going from your PREIND forcings to the 1850 forcings; presumably it might have been a good idea to hold the model steady at 1850 CMIP5 forcings for a few decades before starting the 1850-2005 simulation?
- HadCM3 runs at about 100 SYPD these days (see Valdes et al, 2017). it might be nice to compare some metrics with them, in a consistent style e.g. Geckler plots (see Figure 2 in Valdes et al, 2017).
- In addition to the Cretaceous simulation with the new model grid, it would be very informative to re-do a preindustrial control with this new grid, to see how much of a role the modern-specific tweaks play

Minor corrections

Line 8: “and” missing.

Line 27: “which” should be “whose”?

Line 150: spell out “phytoplankton” and “zooplankton”.

Figure 2 caption: “The surface of the disks” should be “The surface area of the disks”.

Figure 4: label panels (d) and (e) as CRF and temp anomalies (just so we don’t have to read the caption!).

Figure 19b. Black continental outline is missing for some coastlines.

Line 580. Reference missing.

Figure 21: why suddenly a new map projection?

I was surprised that Jean-Louis Dufresne was not a co-author, given his contributions to the original model?

Erb, M.P., C.S. Jackson, and A.J. Broccoli, 2015: Using Single-Forcing GCM Simulations to Reconstruct and Interpret Quaternary Climate Change. J. Climate, 28, 9746–9767, https://doi.org/10.1175/JCLI-D-15-0329.1

Lord, N. S., Crucifix, M., Lunt, D. J., Thorne, M. C., Bounceur, N., Dowsett, H., O'Brien, C. L., and Ridgwell, A.: Emulation of Long-Term Changes in Global Climate: Application to the Late Pliocene and Future, Climate of the Past, 13, 1539–1571, https://doi.org/10.5194/cp-13-1539-2017, 2017.

Lunt, D. J., Bragg, F., Chan, W.-L., Hutchinson, D. K., Ladant, J.-B., Niezgodzki, I., Steinig, S., Zhang, Z., Zhu, J., Abe-Ouchi, A., de Boer, A. M., Coxall, H. K., Donnadieu, Y., Knorr, G., Langebroek, P. M., Lohmann, G., Poulsen, C. J., Sepulchre, P., Tierney, J., Valdes, P. J., Dunkley Jones, T., Hollis, C. J., Huber, M., and Otto-Bliesner, B. L.: DeepMIP: Model intercomparison of early Eocene climatic optimum (EECO) large-scale climate features and comparison with proxy data. Submitted to Clim. Past.

Marzocchi, A., Lunt, D. J., Flecker, R., Bradshaw, C. D., Farnsworth, A., and Hilgen, F. J.: Orbital control on late Miocene climate and the North African monsoon: insight from an ensemble of sub-precessional simulations, Clim. Past, 11, 1271-1295, doi:10.5194/cp-11-1271-2015, 2015.

Rugenstein, M., Bloch-Johnson, J., Gregory, J., Andrews, T., Mauritsen, T., Li, C., et al (2019). Equilibrium climate sensitivity estimated by equilibrating climate models. *Geophysical Research Letters*, 46. https://doi.org/10.1029/2019GL083898

Simon, D., A. Marzocchi, R. Flecker, D.J. Lunt, F.J. Hilgen, P. Th. Meijer, Quantifying the Mediterranean freshwater budget throughout the late Miocene: New implications for sapropel formation and the Messinian Salinity Crisis, Earth and Planetary Science Letters, 472, 25-37, 2017.

Singarayer, JS & Valdes, PJ 2010, 'High-latitude climate sensitivity to ice-sheet forcing over the last 120 kyr', *Quaternary Science Reviews*, vol. 29, no. 1-2, pp. 43 - 55. https://doi.org/10.1016/j.quascirev.2009.10.011

Valdes, P. J., Armstrong, E., Badger, M. P. S., Bradshaw, C. D., Bragg, F., Crucifix, M., Davies-Barnard, T., Day, J. J., Farnsworth, A., Gordon, C., Hopcroft, P. O., Kennedy, A. T., Lord, N. S., Lunt, D. J., Marzocchi, A., Parry, L. M., Pope, V., Roberts, W. H. G., Stone, E. J., Tourte, G. J. L., and Williams, J. H. T.: The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0, Geosci. Model Dev., 10, 3715-3743, 2017.