We would like to thank the editorial office, executive editor and two anonymous reviewers for their comments on our manuscript. Please find below our point-by-point replies. The editor/reviewers' comments appear in purple, our replies in black, and any text from the submitted manuscript in blue, and any new or modified text in green.

Editorial office

1. With the next revision, please do not forget to add the DOI https://doi.org/10.5281/zenodo.14721184 to the "in-text" citation (Tsigaridis et al., 2025) in the section "Data availability" of the manuscript.

This is already listed there. No changes made.

2. With the next revision, please re-name appendix tables: from "Table A4.1" to "Table C1", from "Table A4.2" to "Table C2" and please consider adding captions to other appendix tables. See more: https://www.geoscientific-model-development.net/submission.html#manuscriptcomposition > 7.

Done.

3. Your reference list includes works "in preparation". Such works can be cited upon submission if being available to the reviewers. They should not be cited in the final, accepted manuscript, unless published, accepted for publication, or available as preprint with a DOI.

Unfortunately the two "in preparation" papers will not be ready on time, so we removed them, and replaced them with some text in Sect. 2.2.1: "the new physics will be described in detail elsewhere in the future"; "A publication about these evaluations is forthcoming". We simply deleted the citation in Sect 4.4.3.

4. Please ensure that the colour schemes used in your maps and charts allow readers with colour vision deficiencies to correctly interpret your findings. Please check your figures using the Coblis – Color Blindness Simulator (https://www.color-blindness.com/coblis-color-blindness-simulator/) and revise the colour schemes accordingly with the next file upload request. -> Fig. 7.

We replaced the red color with purple in Fig. 7, and adjusted the legend correspondingly.

Executive editor

Unfortunately, after checking your manuscript, it has come to our attention that it does not comply with our "Code and Data Policy".

https://www.geoscientific-model-development.net/policies/code_and_data_policy.html

You have archived the ROCKE-3D code in a site that does not comply with our code policy. Therefore, please publish your code in one of the appropriate repositories according to our policy.

In this way, you must reply to this comment with the link to the new repository used in your manuscript, and its permanent identifier (e.g. DOI). The reply and the repository should be available as soon as possible, and before the Discussions stage is closed, to be sure that anyone has access to it for review purposes.

Please, note that if you do not fix this problem, we will have to reject your manuscript for publication in our journal.

Thank you for your comment. The model code is already stored in the zenodo public archive together with our simulation output, which contains a DOI. This is mentioned in lines 95-96 of the submitted manuscript:

"The model code, all output, and the model configurations used here, are available on a zenodo archive (Tsigaridis et al., 2025)"

This is mentioned in the zenodo metadata too:

"modelE2_planet_2.0.tar.gz: The model code. This is identical to the code provided in the official model snapshots website (https://simplex.giss.nasa.gov/snapshots/) under "version 2.0 of ROCKE-3D model""

We do acknowledge though that we failed to report this in the code availability section of the manuscript, which only points to our long-term institutional repository that has no DOI associated with it. In the revised version we included a statement about the zenodo repository: "The model code, all output, and the model configurations used here, are available on a zenodo archive (Tsigaridis et al., 2025)."

Reviewer #1

The manuscript describes the new version of the ROCKE-3D model (Planet 2.0). The introduction describes relevant historical developments, before the manuscript moves on to describing model components and some sample simulations made with the model to tune for different scenarios and different combinations of model components. There is also some exploration of model responses from a scientific point of view and differences with the previous ROCKE-3D model version are described.

I think that almost everything that is needed is there and that the paper is quite close to publishable in the journal following the consideration of my minor (but quite lengthy) comments.

Thank you.

General comments:

Given that this is a general circulation model description paper, I was surprised that there was no mention of the model's dynamical core that solves resolved physical equations. Granted this core is taken from the GISS Model E base model, but the same is true for many other model components, which are at least mentioned briefly.

We have added a paragraph describing the dynamical core at the beginning of Sect. 2.

There are unfortunately a rather a large number of minor errors in the paper. I have tried to catch as many as I can, but I think the paper would benefit from the authors taking care to go through it again for readability. With luck, this paper should be useful to modellers for many years!

We apologize for the errors. We hope we captured most of them now, with the help of your comments.

Minor comments:

Line 21: The IPCC is the Intergovernmental Panel ON Climate Change.

Fixed.

L 27: Extra space after "SOCRATES"?

Fixed.

L 31: Think it would be good to say what the atmosphere used in ROCKE-3D 1.0 for readers that don't know what this is.

The exact description is mentioned in the manuscript and it is a little too verbose for the abstract, so we added “the aerosol and ozone-free atmosphere used in ROCKE-3D 1.0” here, since these two are the most important characteristics of that atmosphere.

L 52: Dynamic oceans also include further processes that are not parameterised, but solved on the model grid -- specifically fluid dynamics of the ocean.

This is correct, and we modified the sentence to convey this message: “…that explicitly calculate horizontal and vertical ocean heat transports at resolved spatial scales while also including a range of parameterized processes that operate on smaller unresolved scales”.

L 57: Why are the modellers unwilling to use dynamic oceans? Presumably computational expense?

We added “due to the increased computational cost and additional free parameters, given the lack of information about oceans in exoplanets”.

L 61: I would include units T = -21.

Added.

L 65: I don’t think “demonstrated” is the right word here, as you not actually shown this to be the case in the introduction? Recommend just DELETE “As demonstrated… introduction,”.

Deleted.

L 112: Does the sensitivity of cloud formation to boundary layer height affect cloud outside the boundary layer as well as inside?

We apologize for the confusion caused, this was not properly worded. We modified the sentence as follows: “Another change makes the threshold relative humidity for subgrid cloud formation aware of the planetary boundary layer height as simulated by the model, rather than assuming a fixed pressure altitude that was the case in older model versions”.

L 118: Is the removal of Rossby radius scaling a good thing for planetary science given that ROCKE-3D will be used for planets with different rotation rates? (I have no idea but would like to know!)

We are not recommending the use of the baroclinicity-scaled diffusivity for non-Earth runs. For non-Earth runs, many factors determine whether parameterized mesoscale transports play an important role in the solution, and we instead recommend varying the prescribed diffusivity for the potential exploration of such questions. We now caution the reader for this: “…dependence, an option that should only be used for Earth simulations”.

L 122: Are you saying that the supply of groundwater is unlimited?

We enriched this sentence: “The model has the capability to include irrigation, in which water is withdrawn from lakes and then from an unlimited groundwater pool if lakes are insufficient. This groundwater pool is not dynamic with the full hydrological cycle and does not get recharged from runoff or infiltration (Sect. 2.4.1 in Kelley et al., 2020), so there are small increases or decreases in total planet water mass and sea level depending on whether a particular climate simulation draws from or adds to the groundwater supply. For simulations without humans (i.e. modern Earth), irrigation is not included."

L 146: The prognostic rivers are the ones where rivers are NOT prescribed I assume? Don't think this was absolutely bolted down where the new river routing scheme is introduced.

It is the river direction that is prescribed (or not), which is mentioned in line 143 of the submitted manuscript. Whether a river would ultimately be created is always prognostic in the model depending on the hydrological state at any given time.

L 150: "were not used in this work". I think you mean that they are not explored in the results shown in the manuscript. They are available in ROCKE-3D?

Yes, this is exactly what the rest of the sentence is saying already (in italics here): "which were not used in this work *but are available in the public release of planet_2.0*".

L 157: I think there may be confusion over the use of the term "radiative forcing" here. In the context of Earth climate change, radiative forcing is the change in top of atmosphere radiative flux compared with background values due to the influence of an external factor, such as greenhouse gases, over the course of the scenario being considered. Hence comparison of a background heat flux (I think the Davies value is a mean geothermal flux with no changes predicted?) to radiative forcing isn't quite right. For example, the background solar flux for Earth is about 300 Wm^-2, which obviously can't be ignored, but the solar forcing for the 21st century is expected to be fairly small compared with anthropogenic greenhouse gas forcing. Perhaps a comparison between background fluxes would be more relevant here?

This was a mistake from our side, thanks for catching it. We replaced "radiative forcing" with "energy".

L 204: Is it possible to say briefly why non-LTE corrections are not always required to assist model users in understanding the validity of the model?

We modified the wording and extended the discussion about this: "For exceedingly thin atmospheres (< 10 microbars) non-LTE radiative effects may become important, however, non-LTE physics are not implemented in SOCRATES. For Early Moon simulations we assume CO-dominated atmospheres, in which, fortuitously, non-LTE effects are negligible and can be ignored One should notice though, that in some cases non-LTE corrections may not be needed even for atmospheres thinner than 10 microbar, as is the case for a CO atmosphere (Forget et al., 2017). However, caution should be exercised that for different atmospheric compositions (e.g. CO2-dominated) non-LTE effects can be important and can have non-negligible impacts on results".

Also – "A thin CO atmosphere…" This sentence does not seem related to the rest of the paragraph. What is the relevance of this?

This was misplaced. We moved the sentence at the end of the previous paragraph: "In the current release, the model can handle surface atmospheric pressures down to 10 microbar. This configuration was already used to simulate a thin CO atmosphere as a likely candidate for a transient volcanically-induced paleo-lunar atmosphere (Aleinov et al., 2019; Needham and Kring, 2017)".

L 234: Should "apsis" be "apoapsis"?

Thank you for catching it. You are right, apsis is the generic term for both periapsis and apoapsis. Fixed.

L 245: I think SOCRATES has moved on a long way since Edwards and Slingo (1996). However, not sure I can find an up-to-date reference in a journal. Perhaps cite something like

Walters, D., Boutle, I., Brooks, M., Melvin, T., Stratton, R., Vosper, S., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S., Mittermaier, M., Morcrette, C., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P. L., Wilkinson, J., Willett, M., Woolnough, S., and Xavier, P.: The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, Geosci. Model Dev., 10, 1487–1520, https://doi.org/10.5194/gmd-10-1487-2017, 2017.

Which describes some improvements (although it's almost 8 years old already).

We added two more recent papers that are more appropriate than the one proposed (Amundsen et al., 2016, 2017). Also note that the ISCA model cites the following, but the link is behind the MetOffice wall, so we decided to not cite it here: "Manners, J. and Edwards, J. M. and Hill, P. and Thelen, J.-C., 2015: SOCRATES (Suite Of Community RAdiative Transfer codes based on Edwards and Slingo) Technical Guide. Met Office, UK. Available at: https://code.metoffice.gov.uk/trac/socrates".

L 290: I think "small fraction" might give the uninformed reader the idea that oceanic transport is less important than perhaps it is? I agree that clearly the atmosphere dominates, but that is not to say that oceanic transport is unimportant. There is a clear handing over of heat transport to the atmosphere in the mid-latitudes, but the way the paragraph is written overemphasises the role of the atmosphere (and underestimates the potential role of the ocean in other climates).

We made a significant restructuring of section 2.5, which we invite you to look in the revised manuscript, to avoid repeating a large block of text here.

L 347: Sorry, I don't understand this sentence. "A secondary goal… not relevant to this work." How can something be a secondary goal, but also not relevant?

We rephrased this sentence: "For Earth-centric science questions, a secondary goal is that the present-day net radiative balance of the planet is on the order of +1 W m-2, but this is not relevant to this work which focuses on a generalized planetary configuration rather than the current state of the planet and the impact humans have on it".

L 350: "the atmosphere might have a hard time to adjust." The language is a bit colloquial here I'm afraid. What do you mean specifically?

We rephrased this: "the atmosphere might take much longer, or even fail, to fully adjust".

Paragraph starting L 364: Is the comparison here between planet_2.0 and the base GISS Model E or planet_1.0 or something else?

This is against ROCKE-3D 2.0, since we refer to section 2.2 here. We made that explicit: "planetary (ROCKE-3D 2.0; Sect. 2.2)".

L 379: “disabling of the gravity… with the simpler Rayleigh…” Do you mean “replacement” rather than “disabling”?

We changed “disabling” with “replacement”.

Table 2 and Figure 3. The model versions appear in different orders in the figure and table. Consider describing them in the same order to help the reader?

Done.

L 435: Consider rephrasing. “Throughout the tropics…” suggests that tropical cloud cover has no effect outside the tropics. Not sure that’s what you mean?

We rephrased the sentence: “Cloud cover increases throughout the tropics by as much as 60% locally, but changes less than 5% anywhere else (Fig. 4)”.

L 437: What is the Earth-centric adjustment?

We clarified this: “Earth-centric adjustments (P2GApM40_notEarth in Table 2)”.

Figure 4: The individual figure titles are very small and faint and therefore difficult to read. Please improve. Also, what do the dots on this figure signify? I don’t think they are just a plotting artifact, as they seem to be different on different panels.

We added the explanation of the dots in the figure caption: “Dots show grid boxes where the calculated differences are not statistically significant ($p>0.05$)”. We also increased the dots size, as also requested by reviewer #2.

L 467: “Since no clouds are expected to exist in that model configuration”. Sorry I do not understand this? What simulation are you referring to that has no clouds in it? As I understand it the simulations in table 2 are Earth-like enough that we might reasonably expect some clouds? Perhaps more explanation is needed.

Apologies, this is not what we meant, we should have said: “since no clouds are expected to exist above 50 hPa”, which is the configuration of the model with the cloud-top limit removed. We fixed that.

L 482: “+1.0 (M40) and +1.3 (F40)” are these in Wm^-2?

Yes; added.

L 487: “A key conclusion is the confirmation that virtually all configurations across model versions produce the same climatology. These are marked with a blue arrow in Fig. 5, where there are practically no differences between the simulations.” This statement does not appear to be supported by the figure? There are more than 30 perturbed simulations in Figure 5, but only two of them have blue arrows.

There are three blue arrows, not two, but the important thing is that we understand why the confusion: we speak about model versions, which are four (mentioned by the orange and blue arrows), not model configurations. We enhanced the statement as follows: “A key conclusion is the confirmation that virtually all Earth configurations across model versions (ModelE2.1 (E2.1GApF40_1850), ModelE2.1 with updates (T3GApF40_1850), ROCKE-3D 2.0 starting point (P2GApF40_1850), and final ROCKE-3D 2.0 generalized configuration (P2GApF40); also see Table 2) produce the same climatology”.

L 509: “The simulations do not have a stratosphere”. I had a look into this and I see that you are correct that the Earth literature often defines the stratosphere as a region in which temperatures increase with height. (Even the American Meteorological Society glossary states this!) There is an alternative definition that states that the stratosphere is a region of the atmosphere that is close to radiative balance. This occurs in the absence and presence of ozone and does not require that the atmospheric temperature gradient dT/dz is positive. Only that it is stable. See for example the discussion in the introduction to:

Thuburn, J., and G. C. Craig, 2000: Stratospheric Influence on Tropopause Height: The Radiative Constraint. J. Atmos. Sci., 57, 17–28, https://doi.org/10.1175/1520-0469(2000)057<0017:SIOTHT>2.0.CO;2.

For the present paper I think that this second radiative-balance definition is interesting because it applies across a much wider range of planets than just an Earth-like planet with an ozone layer and highlights the dynamical differences across an atmosphere. Mars, for example, can be said to have a robust stratosphere under this definition. In any case, I think it would be a good idea to state up front what you mean, as initially I was confused. I thought you meant that no part of the relevant atmospheres is close to radiative equilibrium, but analysis of your figures suggests to me that that is not the case.

Thank you for this important comment. We now use the term “stratospheric temperature inversion” instead of “stratosphere” throughout (lines 463, 509, and 977 of the originally submitted manuscript).

L 531: I think you are suggesting that a 20 year mean is equivalent to a 20 member model ensemble mean? I’m not really sure what you are driving at. This sentence could be deleted.

Deleted.

Figure 6: The outliers are individual outlier years or something else?

Yes, individual outlier years. We replaced “outliers” with “outlier years”.

L 565: When you say “is in all climate-relevant respects the same”, do you mean that it produces the same climate or that the code is the same? (I thought the former, but then the next sentence made me reconsider.)

Yes, we mean the different code bases produce the same climate. We appended the last sentence of this paragraph to make it clear: “the model skill and mean climatology do not change with the updated code in ROCKE-3D 2.0 when compared with ModelE2.1”.

Figure 7: x-axis units are years?

Yes. We modified the legend.

L 627: “dependents” -> “depends”

Fixed.

L 646: Should the idea be that the configuration changes are as limited AS POSSIBLE (to reduce spinup time) rather than that configuration changes MUST be limited?

Actually, neither. What we mean here is that for programming reasons a restart file might become incompatible with a model configuration when code changes require things from the restart file that don’t exist there. We have some flexibility coded in for common use cases, but these are in no way capture all possible things a user might do. We modified that sentence: "The key thing to keep in mind is that the configuration changes between a restart file and a new simulation that the model code allows are limited".

L 669: Wasn't the supercontinent called Pangea? (Or Pangaea or maybe the Pangean Supercontinent…?)

We changed this to "the Pangean supercontinent", a term we were using later in the manuscript too.

L 699: The land at the south pole point in the Venus aquaplanet simulation. Is this incorporated for computational reasons rather than scientific reasons?

Yes, and this is the case for all aquaplanet configurations in ROCKE-3D, not just Venus.

L 711 and Figure 10: In the Earth literature at least, the term "cloud radiative forcing" is deprecated. "Cloud radiative effect" is preferred. This is because "radiative forcings" are taken to imply external influences that are applied to the model (such as a change in the composition of an atmospheric absorber, or a change in incoming stellar radiation).

Fixed.

L 779: The "bathtub ocean" has a maximum depth of nearly 3800 meters. Is this a constant depth or is there variation?

We added more information: "include simple "bathtub" oceans, in which the sea floor slopes steeply downward from the continental edges to a uniform ocean bottom. The maximum depth is configurable, and is typically set between 1000 and 3800 meters for simulations with a dynamic ocean".

L 789: To clarify: The runs have global mean surface temperatures of 20.9 and 22.2 C? Or they are 20.9 and 22.2 C warmer than runs with Earth topography?

The former, which can be seen in figure 14 just below that sentence.

L 841: "40x time oceanic increase" Is this a typo?

We deleted the word "time".

L 860: "A comprehensive list and explanation of all options is far outside the scope of this work…" Well – this is the model description paper, so we might expect to see some of the major parameters described? As it is, most of the parameters described relate to timesteps, diagnostics and ice and low temperatures. Given that other model parameters were tuned between various set-ups that will be released to the public, it might be good to describe these since that will be useful not only for understanding differences between simulations but also for others who want to do some model tuning themselves? Looks like there may be some description under "variable names" in Appendix D, but this isn't quite the same thing.

We have added a new Table 4 (the old one moved to Table 5) in Section 4.4 which describes parameters related to a planetary configuration. Going a step further and documenting everything in this manuscript, however, is not a simple task. A comprehensive list of every parameter would be in the order of several hundred, many of which are defined in Way et al. (2017). A description of many of these exist in relevant papers, mostly of Earth focus, or our public documentation, but not all. The important thing to note is that we purposefully avoided mentioning too many parameters that one can modify, since that might do more harm than good if someone tries to make a change without properly understanding the model mechanics. A good example is the flag use_vmp, mentioned in line 1010 of the submitted manuscript, as well as in Table 4 and Appendix B. We cautioned the readers that if one decides to change this, a rebalancing would be necessary. Another example is NRAD, extensively described in lines 869-880 and in Table 4, incorrect use of which can significantly bias the model output, resulting in erroneous results, especially for slow rotators. Many (maybe even most) of the parameters one can use might have unwanted side effects, which even experienced model developers might have a hard time disentangling. For that reason, we presented a rather long list of parameters that based on our experience with ROCKE-3D over the years a user (not a developer) should be aware of. For anything more, again based on our experience with the public release of the previous model version, people do reach out to us and we help them as needed. Also note the outreach section of the submitted manuscript, which mentions the annual tutorials we are holding for new users.

L 863 and table 4: In Line 863, DTsrc is described as the parameter that "tells the model how often its main parts should exchange fluxes" – so the frequency at which coupling between atmosphere and ocean etc occur. However, in table 4, DTsrc is described as the physics timestep (presumably atmospheric processes that are not simulated by the dynamical core?) Is one of these definitions in error?

Yes, the description in Table 4 was incorrect, thanks for catching it. It now reads: "Timestep for exchange fluxes across model components".

Table 4: maxctop "can affect model results". Does this mean that the parameter occurs in the prognostic equations. I recommend that this is stated given that the previous two parameters are given as diagnostic.

We added the term prognostic: "can affect prognostic model results".

L 902: "The first thing one notices when comparing Fig. 5 with Fig. 16 is how much more colorful the latter is, implying that the simulated differences across planetary configurations are much more impactful to climate variables than the changes we did towards the generalization of the Earth configuration." I'm not sure I follow I'm afraid. I think you are saying that the differences between planetary configurations are larger than the differences between GCM setups for a given planetary configuration?

Yes, with the note that "a given planetary configuration" only means Earth, since we can't really have planet-specific tunings for any other planet due to lack of data. We added a sentence to make our statement clearer: "In other words, the climatological differences we get when simulating Earth with and without any Earth-centric tunings is much smaller than the differences seen when simulating different planets".

L 922: Just a check: Is this the tropopause temperature and pressure or that of the temperature inversion?

We explained what we mean here, by adding a sentence at the end of Sect. 5.1: "ROCKE-3D uses the definition by the World Meteorological Organization (WMO) to calculate the tropopause layer, as the lowest level at which the temperature lapse rate decreases to 2 K km-1 or less".

L 976: Another reference to the stratosphere. Would be good to revisit the definition.

Please see reply to comment L 509.

L 1009: "Give way abruptly to" or "abruptly give way to".

Fixed, we picked the second recommendation.

L 1035: Heat that is "deposited near the tropics" because of the star or something else?

This is because of Earth's obliquity. We clarified that in the text.

Figures 24 and 25: Figure captions / legends are faint here. Would be good if they could be made bold.

We increased the thickness of all lines and fonts in the two figures.

L 1116: "Even with this numerical noise… regardless." I'm not sure this sentence makes sense? Please check.

We rephrased the sentence: "This numerical noise though is random and does not introduce biases, which means that the climatologies calculated by any computer are virtually identical after long integrations. Long integrations are needed regardless for a proper analysis of results, so a change of computer is not expected to affect results qualitatively, and even quantitatively any change will not be statistically significant".

L 1118: "quantitatively the same". Do you mean "qualitatively the same" given that values are not – as you say – identical?

Yes, thanks for catching this. The rephrasing mentioned in the previous question should give more clarity on that statement too.

Conclusion section: These are worthy aims. However, it would be good if the conclusion section also provided a short summary of the findings of the paper, perhaps with some forward look. I think this is particularly true given that the introduction takes the time to give a nice summary of relevant past work on numerical modelling. What are the most important avenues for future model development?

We have greatly enhanced the conclusions section by including both a summary and a list of future development goals, which is now four paragraphs, instead of one in the submitted version.

Reviewer #2

The paper describes the second version of ROCKE-3D, which includes upgrades derived from both Earth-centric developments (such as physics updates, bug fixes, and structural enhancements from the GISS ModelE2.1) and planetary-centric developments (notably in radiative transfer calculations, atmospheric composition, ocean parameterizations, and model resolution). It also introduces specific options tailored to the modeling of rocky planets, including thin atmospheres, geothermal heat flux, and an updated calendar module.

The paper represents a valuable contribution and a substantial advancement in the modeling of rocky planets. However, the structure should be improved, and some simulation results need a more in-depth discussion, as outlined below (with additional suggestions included in the list of minor comments).

Thank you.

First, for completeness, a brief description of the atmosphere, sea ice, continental ice, and vegetation components used in ROCKE-3D should be provided in Section 2.

We greatly expanded Section 2 by over a page, which now describes the core components of the model, including its dynamical core.

As shown on p. 33, when using the SOCRATES radiation scheme with Qflux = 0 at 1xCO_2, the simulation results in a snowball state. This implies that the SOCRATES runs at 1xCO_2 and 2xCO_2, starting from the same initial conditions, lie below and above the unstable branch that separates the warm and snowball states, respectively. This behavior is both interesting and expected, not only in slab-ocean models but also in GCMs (e.g., Ferreira et al., J. Clim. 24, 992 (2011); Brunetti & Ragon, PRE 107, 054214 (2023)). The fact that the GISS radiation scheme reaches the snowball state at a different CO_2 concentration (0.5xCO_2) simply indicates that the bifurcation diagrams for SOCRATES and GISS differ—particularly regarding the location of the unstable branches.

A similar behavior is expected when using a dynamic ocean. Therefore, it would be useful to expand the discussion at this point and, if possible, include additional simulations showing the occurrence of the snowball state also in the case with a dynamic ocean.

We agree that probing the bifurcation/hysteresis structure of the many ROCKE-3D configurations would be a valuable exercise, although as a model description paper we believe this is best left for future work. This is both because it is an interesting scientific question on its own merit (particularly how the Q-flux and dynamic oceans affect the snowball transition or deglaciation), but also because it would greatly expand the number of simulations beyond what we intend to present for the current work. We highlighted the particular case of the SOCRATES/GISS Q-flux ocean since it is so close to an instability, but we saw no evidence the dynamic ocean configurations are similarly close to instable conditions. You can also refer to our past publication, Colose et al. (2019; doi:10.3847/1538-4357/ab4131), regarding the hysteresis structure of ROCKE-3D v1.0.

As you mention in lines 802–804, sea ice was removed in Rencurrel & Rose (2018) because their study specifically focused on ocean–atmosphere interactions. However, more generally—and depending, of course, on the scientific question being addressed—sea ice should not be excluded from the model. Its role in the full coupling between atmosphere and ocean is crucial for climate dynamics, particularly for the emergence of steady-state climates characterized by large ice caps, waterbelt configurations, or snowball conditions (Rose, J. Geophys. Res. Atmos. 120, 1404 (2015), Brunetti et al., Clim Dyn. 53, 6293 (2019), Hörner & Voigt, Earth Syst. Dyn. 15, 215–223 (2024)), all of which are highly relevant in exoplanet studies. The discussion in this section would benefit from being expanded and refined to better reflect the importance of sea ice in such contexts.

We agree that sea-ice should not be excluded in modeling studies where standing water is simulated, either in the form of lakes or oceans. We added the following: "While sea ice is sometimes disabled in some aquaplanet studies to isolate specific dynamical processes, its inclusion is essential for fully coupled exoplanet simulations. Sea ice plays a critical role in setting the climate state, particularly in transitions between ice-free, waterbelt, and snowball regimes and tipping points, and strongly influences the planet's energy balance and climate stability (e.g. Brunetti et al., 2019; Hörner and Voigt, 2024). For the realistic exploration of exoplanet climates, especially near habitability boundaries, sea ice must be actively included".

Figure 18 is particularly interesting for understanding how key mechanisms depend on the choice of radiation schemes in dynamic ocean simulations. The GISS scheme produces a warm blob in the northern hemisphere when using atmosphere A, associated with reduced total cloud cover. This result is somewhat counterintuitive and may be linked to processes occurring in the stratosphere. Could the warm blob also be influenced by ocean circulation? The paper does not discuss the strength of the overturning circulation, and including such results would be valuable.

When considering atmospheres x and N, the warm blob disappears, suggesting that it may be related to cloud physics—although it is difficult to draw firm conclusions, given that simulations A, x and N are equilibrated using different parameters, as shown in Table 3.

We added an explanation of this in Sect. 5.2: "The behavior observed in the A-atmosphere configuration appears to stem from the absence of an active overturning circulation in the SOCRATES radiation scheme case for the M-resolution, which effectively shuts down the northward transport of heat. In contrast, the simulation using the GISS radiation scheme maintains a robust Atlantic Meridional Overturning Circulation (AMOC) of approximately 23 Sv, accounting for the warmer temperatures over the North Atlantic and the associated reduction in sea ice area relative to the SOCRATES run. While the radiation scheme does not directly control ocean circulation, its influence on surface energy fluxes may indirectly affect AMOC stability. In our model, AMOC strength has been shown to be highly sensitive to freshwater input from ice melt and transport (Romanou et al., 2023), suggesting a possible link. We also note that at the higher F-resolution, both the SOCRATES and GISS simulations exhibit an active AMOC".

In the x simulations (central column in Fig. 18), the southern hemisphere becomes warmer with GISS and features increased cloud cover. However, the northern hemisphere becomes colder—also with more clouds. Could you provide an explanation for this behavior?

We added the following explanation in Sect. 5.2: "In the x simulations (central column in Fig. 18), the southern hemisphere becomes warmer with the GISS radiation and features increased cloud cover. However, the northern hemisphere becomes colder, also with more clouds. This particular GISS-SOCRATES difference appears to be the superposition of two other GISS-SOCRATES differences: one, the model wants to be cooler at high latitudes with the GISS radiation, excepting the North Atlantic AMOC-related feature, and two, the model wants to be warmer at southern high latitudes when using the GISS radiation (and slightly warmer at northern high latitudes) and when changing the atmospheric composition (and cloud tuning) but holding the ocean circulation fixed. When we combine these two patterns, we get the southern hemisphere warming in Fig. 19 slightly winning out over its cooling in the upper left of Fig. 18, but in the northern hemisphere the Fig. 18 cooling wins over the Fig. 19 warming". Note also that we moved part of the pre-existing discussion right after Fig. 17, rather than before, to be closer to Figures 18 and 19.

In the N simulations (right column in Fig. 18), the equatorial regions become warmer and less cloudy with GISS. Similarly, in Figure 19, the qN simulations show that GISS yields a warmer equatorial region with reduced cloudiness. These findings raise important questions about the cloud formation mechanisms at play, which would benefit from a deeper discussion.

Although it is a challenging task to provide concise and precise explanations across model configurations with different cloud tuning parameters, we tried to provide some reasoning about the differences seen in Figures 18 and 19. This is particularly true when trying to balance not-quite-Earth configurations (i.e. the x and N atmospheres) using Earth SSTs, where a lot of subtle things can happen. We added the following discussion right before Fig. 19: "The differences in cloud cover between the GISS and SOCRATES simulations are generally consistent with the differences in the cloud tuning parameters, at least in the extratropics where there are the most stratiform clouds (the tuning parameters are for stratiform clouds). In particular, the A-to-x changes to those parameters for the GISS radiation are very large compared to typical changes for rebalancing Earth, and go in the direction of making Gx a cloudier world. Interestingly, the SOCRATES radiation did not require such a large change for A-to-x rebalancing, perhaps because its *radiusl_multiplier* was left unchanged, while the A-to-x rebalancing for the G radiation increased *radiusl_multiplier*, which tends to make clouds less opaque and thus requires increases of cloud cover to compensate (decreased *U00a*, increased *U00b*)". To thoroughly explore these differences one should conduct experiments where not-Earth atmospheres using Earth SSTs are not rebalanced, and just let the q and o models find their new balanced equilibria for both Gx/Sx and GN/SN pairs.

Minor comments and corrections:

l30, Abstract: specify which is the atmosphere used in ROCKE-3D 1.0

Done. Also see our reply to reviewer #1 on the same point (L 31).

l54: the sentence is not clear in my opinion, it seems that slab-ocean (Q-flux) can be considered AOGCM while they do not have a dynamical ocean (thus, not OGCM by definition). The sentence needs to be reformulated.

We replaced "AOGCMs" with "a parameterized ocean" and further rephrased the couple sentences after it, following also a comment from reviewer #1 (L57).

Section 2, Model description: I suggest to use only ROCKE-3D 1.0 and 2.0, instead of adding also planet_1.0 and planet_2.0, which are redundant (and confusing) definitions of the same models. (Note that at l. 98 another name is used: ROCKE-3D Planet 1.0).

Done.

l83: `in section 2.1' instead of `in the same section'

Fixed.

l122: `small increases and decreases' : how much water goes in groudwater recharge? How well is the water mass conserved?

We enriched this discussion, please see the answer to reviewer #1 for the same line.

l130-131: use \citet for Rosenzweig, Russell and Schmidt refs.

Fixed.

Section 2.2.1: Dynamical lakes are tested elsewhere. However, it would be useful to know which is the impact of including them, in particular at which resolution one needs to include their dynamics.

We added a paragraph: "The processes that accumulate water in lakes, combined with the horizontal transport of water across the land surface via runoff and river flow, are important for capturing the mass water balance on land. The geographic distribution of lakes determines the evaporative surface area, the surface cover of liquid water (which can produce glint as a sign of liquid water), the regional formation of clouds, the surface energy balance, and temperature. Previous ROCKE-3D 1.0 experiments could have lakes either without river transport, leading to potential excess buildup of water in some grid cells, or with prescribed river directions in a single direction, which could still result in excess buildup and generally lacked realism in the spread of water across the land surface. Fully prognostic, multi-directional river transport reproduced the retention of water in areas that are in fact wetlands, promoted continental recycling of precipitation, and significantly rectified the warm bias. The generalized lake/river dynamics also enable the simulation of surface hydrology of idealized planets for which only a flat topography is specified, in which case river directions cannot be estimated from topography but must arise from the relative balance of surface water between grid cells."

l184: millibar is sometimes denoted as mb and other times mbar (the second is prefereed, however should be consistent everywhere)

Fixed in 5 places, mbar is used throughout now.

l189: `stellar' instead of `solar' (check also elsewhere in the text)

Fixed throughout the manuscript, where applicable.

l210, 213, 222…: `integer' instead of 'integral'

Fixed.

l258: speed reduction: you can point the reader to Sec. 4.5 and Fig. 15

Done, in the previous sentence.

Sec. 2.5: As you discuss at p. 33, it is possible to simulate aquaplanets with flat bathymetry and full dynamical ocean. You should specify this here and point to the paragraph at p. 33.

We now mention this at the end of the section: "The last configuration ("dynamic ocean"; denoted as "o") is a fully dynamic ocean, similar to what was used in ROCKE-3D 1.0, with present-day Earth's bathymetry, but any configuration can be used, e.g. a bathtub ocean with flat bathymetry used in aquaplanet simulations (Sect. 4.4.4)".

l339: `In this work, the GCM variable names of the model output are in italics for many quantities…'

Fixed.

Sect. 3.2: specify that these simulations are performed with prescribed temperature (p).

We added "…by performing 30-year simulations with the prescribed ocean (p) model".

l369: sect. 3.1 instead of 4.1

Fixed.

l384: Can you specify which is the `geographic adjustment' needed in the Earth-oriented version?

We replaced this sentence with a more explanatory one: "Furthermore, the Earth-oriented version has a special tuning of its "Rayleigh friction" near the poles; this was especially important for getting the correct stratospheric wind structure in M configurations, because they completely lack the parameterized gravity wave momentum transports that shape stratospheric winds in later model generations (F resolution). This geographic adjustment is dropped in non-Earth-oriented simulations".

Sect. 3.3: is the 1850 sim the same as in sect. 3.2?

We believe the reviewer asks about the first sentence of Sect. 3.3. This is simulation E2.1GApF40_1850 (we clarified that), which is mentioned in Table 2 in the same section. The 1850 simulations mentioned in Sect 3.2 are more than one, one of which is this one.

Fig. 3 caption: `explained in Sect. 3' instead of 4

Fixed.

l425: `sect. 3.1' instead of 4.1

Fixed.

l437: which are the `Earth-centric adjustments' that you mention here? It is not clear if you refer to irrigation, etc mentioned at the beginning of this section, or to cloud parameterizations mentioned later, at p. 17. Please specify and, in case, reoganise this section accordingly.

We clarified this (also see Reviewer #1 comment L 437).

l440: which is the atmospheric composition in these sims?

We clarified this: "…the choice of atmospheric composition described here (no ozone, aerosols, and stratospheric water vapor from methane oxidation; x configuration in Table 1)".

l467-470: These sentences are not clear, please reformulate.

We rephrased the sentences: "The clouds change (allow them to reach the model top) in the SOCRATES simulation did not change either of the diagnostics beyond noise, since no clouds are expected to exist above 50 hPa. Although no incremental simulation with the clouds allowed to reach the model top was performed using the GISS radiation, it is expected that it will not affect the climatology, as was the case for SOCRATES". Also see comment L 467 from reviewer #1.

l479: which is the longwave correction you mention here?

This is for the noRADN4 case, as described in Table 2. We clarified this in the text.

Sect. 4.1, first paragraph: Give definitions of U00a, U00b and the other two parameters.

All italicized variables are defined in Appendix D, which is also mentioned at the beginning of Sect. 3 (lines 339-341 of the submitted manuscript).

Caption Fig. 6: how long are the simulations with dynamical ocean? 2000 yr as you say later, at p. 23? Specify how long are these runs for obtaining the values in the figure.

This is not the same for every run, and even the original statement of 500 years in the submitted manuscript is not accurate, apologies for that. The legend now says: "…for all ocean configurations following balancing. The last 100 years of 200 to 500-year-long simulations were used for the p and q ocean configurations, which were longer (1000 to 2000 years) for the o ocean".

l566: for regional differences across model simulations (with prescribed sst) you can refer to fig. 5

Figure 5 shows global means, not regional differences. What we would need is something like Fig. 4, but since this sentence says "very minor" we decided to not include a figure. No changes made.

l580: in Table 2 for intermediate simulations, and in Table 3 for the final simulations

Table 2 shows simulations and their description, while Table 3 shows parameters that were used across the final simulations, but does not describe the simulations themselves. We believe it would not be appropriate to combine them in this sentence. No changes made.

l610-611: in the caption of fig. 7 you say that planet 1.0 is shown in green, while in the main text you say it is the faster to equilibrate (while it is the blue)… Please check, since you mixed up the colors in the main text.

Thanks for catching this error; the figure is correct, the text was wrong. We fixed it: "The dynamic ocean in ROCKE-3D 1.0 also equilibrates much faster, in about 1000 years, while the equivalent simulation in ROCKE-3D 2.0 needs over 1500 years. The fine resolution model equilibrates much faster, in about 500 years".

Fig. 8: at which depths levels 11 and 13 correspond? Deeper levels should require more time to stabilise.

We added two sentences in the figure legend: "The mean depth at level 11 is 1129 m and at level 13 is 3868 m" and "It is important to note that deeper oceans take longer to come into equilibrium".

Sec. 4.4.1 and fig. 9: Have these maps for the ancient Earth a constant ocean depth? Which is this depth?

We added a paragraph under Fig. 9: "With the exception of the mid-Pliocene reconstruction, which includes reconstructed bathymetry as outlined by Haywood et al. (2016), other maps of ancient Earth have a uniform ocean depth. This uniform depth is a simplification given that the details of ocean bathymetry are poorly known, if at all, further back in Earth's history. A reasonable depth for ancient Earth simulations is 3800 meters; for "Earth-like" exoplanet simulations using these land distributions and a dynamic ocean, in the absence of any better constraint, we frequently use a depth of 1000 meters".

l692: Say that the modern-day Venusian topography can be found in Fig. 5 of Way & Del Genio 2020

Done.

Fig. 10: I think that this figure is not necessary, since just the final values at the equilibrium are used, which can be put in a table.

We prefer to keep the figure, since it shows not only the equilibrium point, but also the equilibration time, which was analyzed in Figures 7 and 8.

l738: \citet

Fixed.

l740: … and albedo, and conditions at \sim 3.5 Ga.

Fixed.

l748: Here, it is…

Fixed.

l813: this paragraph on `Deep ocean worlds' is describing a future update, I suggest to include it later when you talk about `Other updates', which could become `Other updates and future developments'

Moved.

l823: h/R in math font

Fixed.

l827: parenthesis should be removed

Fixed.

l866: give a definition of DT and Nisurf or point to Table 4

This whole section was referred to Table 4 just 4 lines earlier: "The key ones are described below, and many more are listed in Table 4". No changes made.

l919: you should point the reader to the right variable in Fig. 16 for the low cloud fraction (pcldl?). The same for tropopause temperature and pressure (cldtpp and cldtpt?)

These would be pcldl, ttrop, and ptrop, respectively. Added.

Fig. 17: Add S, G, p, q, o at the top (as for example in Fig. 16) for clarity.

We have updated Figures 6, 17, and 23, based on this suggestion.

Fig. 18: points are not visible, try to increase marker size or type

We have increased the marker size to Fig. 4, Fig. 18, and Fig. 19.

l987: … smoother transition in specific humidity.

Fixed.

l987-988: I would eliminate the last sentence. All this is better explained in the following section.

Deleted.

l1002-1004: Reformulate the sentence: `This is directly …. A atmosphere,…'

We rephrased it: “This is directly explained by the much colder temperatures in the x and N atmospheres at the altitudes where the stratosphere exists in the A atmosphere”.

l1024: I would introduce panel labeling and say: `see, e.g., Fig. 19a in comparison with 19d and 18c for the changes …’

We opted to make it a little more verbose and not use panel labeling (the figures are too busy already), but thanks for the suggestion to directly cite Fig. 19 as well, on top of Fig. 18: “(see, e.g., top-left panel of Fig. 19 for the p ocean vs. the bottom-left panel of the same figure for the q ocean vs. the top-left panel of Fig. 18 for the o ocean, for when looking at the changes due to the radiation scheme selection; other panels in those two figures present additional variables)”.

l1029: (Fig. 18c and 19d), for dynamycal ocean and Q-flux, respectively.

We deleted this reference, since we made the previous statement more explanatory.

l1042: … invalid, we decided …

Fixed.

Fig. 24: last two panels are inverted with respect to the description in the caption.

We fixed the legend to correspond with the figure.

l1111: … in Sec. 4.3 and App. B

Added.

l1023, Outreach: add links

Following the journal’s style guide, the links are presented in the references section as citations. No changes made.

l1046: gridbox^{-1} on the same line

We assumed l1146. Fixed.

l1048: `… 680 mbar. It also avoids…. ‘

We assumed l1148. Fixed.

l1058: add link for ROCKE-3D spectral files

This line appears to be irrelevant with the spectral files, so does line 1158 (if we assume a same typo as earlier). The spectral files are already listed in Appendix C and their location is cited there. No changes made.

Additional changes in the manuscript

Figure 1 legend: “Modern Earth geothermal heat fluxes…”

Following the recommendation of reviewer #1 about figures 24 and 25, we have also improved the line and font thickness in Figures 3, 4, 6, 7, 17, 18, 19, 20, 21, and 22.