Dear Editor,

Please find our responses to the concerns raised by the reviewers with regards to our manuscript,

"*Setting up the physical principles of resilience in a model of the Earth System*".

We believe the that the comments made by the two reviewers have greatly improved our manuscript, and we wish to thank them both for their effort.

Besides our replies to each comment made by the reviewers, we have marked our changes in **bold** in the revised manuscript. Figure 2 has been revised in accordance with comments made by Rev. #2, which has made it more informative.

The reply to the general and specific comments of the referees are the following:

Referee 1:

The authors extend the Earth System model developed in some of their previous works (Barbosa et al., 2020; Bernardini et al., 2025; Bertolami & Francisco, 2018, 2019) by establishing the physical principles underlying resilience features using the Landau-Ginzburg Theory (LGT). Overall, after the first five well-crafted paragraphs of the introduction, I struggle to grasp this manuscript. I suggest the following improvements before it should be considered for publication.

The answers to the specific points are as follows:

Question 1. What is the problem the authors solve with this work? Of course, one can employ various techniques and theories to construct models of almost anything. But what specific gap in our understanding of how the world works or policy problem does this work address?

REPLY: Our main motivation for this work is to capture the physical principles underlying the resilience the Earth System (ES) shows in resisting to change from the Holocene equilibrium state to a new state, whose average temperature is higher, and its sustainability is not ensured. This issue is raised in the question 3 of the referee and it is central to our discussion, but we welcome this question as it allows us to reflect upon the main points of our model.

The LGM of the ES contains its main thermodynamical features and, in its simplest form, shows equilibrium states that can be easily interpreted as the stable Holocene period, a hotter equilibrium state, presumably a Hot-House Earth state, and a colder state that can be associated to a glaciation. Human activities and the higher concentration of greenhouse gases it ensues, can be treated as a perturbation that destabilizes the Holocene and introduces a new and hotter equilibrium state. The model allows for relatively simple calculations and provides straightforward predictions without the need to add unnecessary complexity and to rely on supercomputer simulations. Given the present level of human activities, the model predicts that the Hot-House Earth scenario is an attractor of trajectories, meaning that it is the ultimate outcome of the present Anthropocene evolution. Furthermore, using the Planetary Boundaries it is possible to consider the model as an accounting system of the human activity and to show that it predicts, with accuracy the ocean acidity, from the recent evolution of the CO_2 concentration (Barbosa et al., 2020).

However, the original model did not exhibit features that resemble resilience. Introducing these features is the purpose of the present work. As the model is based on thermodynamical arguments, our goal was to seek for physical properties that would lead to the resilient behavior of the ES. Our results show that resilience is associated to the existence of metastable states and explicit dissipation of energy that prevent the ES to runway towards the Hot-House Earth state. This understanding can now be regarded as a matter of policy as one can think in implementing and boosting these physical properties.

To clarify these points, we have added some text at the introduction of the paper with the above discussion.

"However, the original model did not exhibit explicit features that resemble resilience. This is the main purpose of the present work. As the model is based on thermodynamical arguments, one must seek for physical properties that would lead to a more resilient behaviour of the ES. In the context of the model, resilience is regarded as the resistance the ES shows in changing from one equilibrium state to another. At the present transient period, the Anthropocene, one infers from a multitude of observations that the ES is moving away from the Holocene equilibrium state to a new state, most likely a Hothouse Earth state (Steffen et al. 2018). As we shall see, our results show that resilience is associated to the existence of metastable states and explicit dissipation of energy that prevent the ES to runway towards the Hothouse Earth state."

Question 2. How do the authors define resilience in their paper? The authors often use vague formulations, such as "resilience features", "resilience can be regarded as ...", or "resilience can be associated ...". Does that mean the authors DO regard or associate resilience with the things that come after? What exactly are the resilience features the authors refer to?

REPLY: As stated above, in the context of the LGM, resilience, a broad and complex issue in general, is regarded as the resistance the ES shows in changing from one equilibrium state to another. At the present transitional period, the Anthropocene, one infers from a multitude of observations that the ES moving from the Holocene equilibrium state to a new state, most likely a Hot-House Earth state.

We stressed this point in section II of the paper.

Question 3. What exactly do the authors mean by the use of their term "physical", as in "physical principles", "physical terms", or "physical conditions"? Judging from Eq. 1, it cannot refer much to the actual physics of the Earth system, as in, for example, atmospheric physics.

REPLY: As pointed out above, the original model did not contain features that resemble resilience. The goal was then to introduce them. Given the thermodynamical nature of the model, we were supposed to seek for specific physical properties that would prevent the evolution towards the Hot-House Earth state. We believe to have convincingly shown that this requires the existence of metastable states and explicit dissipation of energy. The model setup provides means to avoid the undesirable evolution by suitable managing of the Planetary boundaries and control of the average temperature. This knowledge can be turned as matter of adaptation and mitigating measures, stewardship policy and geoengineering.

Question 4. The authors write about the "properties that any realistic model of the Earth System should have". In my view, there is no such thing as a REALISTIC model. All models simplify in one way or another. But I am curious about what exactly these properties are.

REPLY: We agree with the referee. The adjectivization is inappropriate in what concerns our model. We have removed the word from the text.

Referee 2. David A. MacKay

Review of "Setting up the physical principles of resilience in a model of the Earth System" (EGUSPHERE-2025-2163) for EGU Earth System Dynamics Paper

Summary:

In this paper, the authors present a way of representing Earth system resilience as metastable states within a simplified thermodynamical model of the Earth system. By representing human impacts with respect to the planetary boundaries, and constructing a stability landscape for the system's free energy, the authors show that self-stabilising attractors can exist for both a Holocene-like and a "Hothouse Earth" state in this model, with the Earth's state being resilient when situated within either of these states.

General comments:

In general this is an interesting paper that helps to formalise some of the proposals for Earth system resilience – and in particular the hypothesis of Holocene and Hothouse climate attractors – within a thermodynamical framework. However, I think the paper can be reorganised somewhat to make the argument and methodology clearer, and there are a number of assumptions that I think should be more explicitly unpacked in order to better contextualise the work.

On structuring, I think the Introduction can be clarified and streamlined to make the background and contribution of your paper clearer. As it stands, it's rather complex in parts, and could do with more clearly introducing the evolution of your approach in previous papers (so it can be broadly understood without having read them), the assumptions made about Earth system states, and what's new here, leaving more technical details to the second section. Indeed, the second section already does more of this, and some of the Introduction could be merged in there. Some of the thermodynamic terminology and theory in the following sections could also be introduced and explained more clearly, such that a non-physicist reading this (being in an inter-disciplinary journal) could get the gist, if not all the technical details, and be able to interpret the results without recourse to prior papers. There are also some parts I think would work better in different sections, as noted in my specific comments below.

On the modelling itself, I like the demonstration of how Earth system attractors can be represented in a simple thermodynamical model, but I think the paper's narrative would work better if you noted early that the possibility of Earth system attractors was inherent to the proposal of e.g. Steffen et al. (2018) in how it described Glacial vs Interglacial vs Hothouse states, and then that you are providing a simple model here of how this can be formalised via thermodynamics with the help of planetary boundaries as a simplified metric of human impact. This framing would also help mitigate some of the concerns I have on making it clearer what simplifications and assumptions are being made in order to simulate the Earth system in this way, and that whether the actual Earth system is approaching a boundary between metastable states remains hypothetical. The PBs, for example, are useful constructs for measuring how Holocene-like the Earth's current state is, but they are not fundamental variables in themselves, and while one can indeed simplify things in a model like this by using them as such, this assumption should be clear. Similarly, while there is some palaeo & model evidence for climate attractors in general (e.g. Westerhold et al. (2020), Brunetti & Ragon 2023, etc.), that Holocene stability indicates we're in an attractor now, and the proposal of Steffen et al. (2018) of a 'Hothouse Earth' state that's accessible with anthropogenic pressures, both remain hypothetical (and the latter particularly speculative), with limited observational evidence in support. It is fine to take these hypotheses as a given of course for a modelling exercise such as this, but – as with using PBs as fundamental variables – that this is being assumed should be explicit, so that this is not taken as proof of their existence, rather as how it might be understood if it is the case.

Specific comments (by line no.):

Abstract – line 4-6t: Is future sustainability and planetary habitability's dependence on Earth system resilience demonstrated? I'm not sure it's discussed in the Introduction. Also, does it logically follow that if sustainable futures depend on ES resilience, then a model to frame it in physical terms is particularly relevant? I think the motivation can be a bit more clearly articulated here.

Abstract l9-10: I get what you mean here, but describing planetary temperature as being dependent on PB interactions is questionable to me – the PBs represent some key ES processes, some of which are indeed important for determining global temperature, but the PBs are not fundamental variables in themselves that can determine climate, such as solar input, albedo, etc., which are all collapsed down in to a simplified radiative forcing anomaly for the climate PB. I'd say "affected by multiple PB interactions" captures these subtleties better.

Abstract l11-12: What does "runaway of the PBs" mean? I can imagine climate runaway, but it's harder to see how e.g. ocean acidification, nitrogen pollution, etc. can runaway. This is further to PBs being most usefully seen as heuristic yardsticks to measure how close the planet is being pushed away from Holocene-like conditions, not fundamental things or processes in themselves, even if they indirectly map to some processes (but to me, it's still ideal to analyse those processes directly when looking at Earth system dynamics, but which you can explicitly justify substituting PBs for in order to make the model tractable).

Abstract l14-16: Is it humanity or the planet moving away from Holocene conditions? I'd say the latter in this context. And conversely, is it the ES or people who can follow sustainable trajectories?

REPLY: These are fair points and we have rewritten the abstract to better specify the issues in question.

Furthermore, two sentences have been added to the Introduction section to clarify the relationship between humanity, the earth system, and desirable/undesirable resilience. **"However, the Holocene epoch has been marked by an unusually stable climate compared to previous geological periods. This has allowed for the development of agriculture, permanent settlements, and the emergence of complex human societies. Hence, from this perspective, a Holocene(-like) state can be deemed desirable, and the safeguarding of resilience of this state of critical importance for humanity."**

L21-23: You spend some time introducing and explaining tipping points in this paragraph, but you don't come back to them much after this – I suggest either integrating this with e.g. more discussion of their proposed role in creating Earth system attractors, as suggested by Steffen et al (2018) and that your model goes on to provide a formalisation of, or discuss later how your results relate to the idea of tipping points. Probably should have a citation for defining TPs here too.

REPLY: Citation is added.

L32-24: Given large uncertainties on tipping cascades, and the still hypothetical nature of Steffen et al.'s proposal, I'd suggest "may" lead to cascading behaviour, and "could" be colossal consequences.

REPLY: Implemented (in bold).

L37: I'd say that resilience theory is ancestral to and closely associated with TP theory from the start, with both sharing roots in dynamical systems theory and its influence on ecology & climatology respectively. Similarly, PBs were framed from the outset with TPs in mind for at least some boundaries.

L43: It's just the "safe operating space".

REPLY: Implemented.

L52-54: I think this could be unpacked some more – why do they link up? You could e.g. discuss (& possibly show with an illustration) how resilience can be mapped to the negative/dampening feedbacks in an attractor, while going over the attractor's basin boundary can lead to tipping to a different attractor/state, which'd help a less familiar reader understand later results. Furthermore, probably should clarify that a system can have resilience without tipping points being present too, as they don't exist for all complex systems, and the link from resilience theory to PBs could also do with a bit more explicit unpacking.

REPLY: This is implemented in the new version of the text. Our new Figure 2 will hopefully convey this message also to the general reader.

L62-65: It needs to be clearly outlined at some point that the existence of these alternative stable climate states is being assumed here, as while there is evidence for climate attractors in general, the presence of ones for e.g. glacial/interglacial/beyond1.5-2C remain hypothetical.

REPLY: This is clarified in the new version of the text.

L69-72: This definition may need grouping with next section, as is perhaps a bit technical for an Introduction and I think would fit flow better there.

REPLY: This is implemented in the new version of the text.

L79-80: This is a good and clear advance on previous work, connecting with the socio-ecological framings mentioned – this would work well along with a clearer unpacking of what previous work did.

REPLY: This is clarified in the new version of the text.

L91-92: Are these two trajectories identified part of your paper's background (i.e. from past studies), method (assumptions you're making) or results? It's unclear in this paragraph which.

REPLY: This is clarified in the new version of the text.

L105-106: The background for this paper might be clearer if these papers were described in turn (in broad terms, leaving the mathematical description of what each did to the next section), thereby introducing your approach and assumptions, and making the novelty of how this paper builds on them clear.

REPLY: This is implemented in the new version of the text.

L109-110: Only this state? Going to a colder glacial one is unlikely given current trajectory, but technically possible, and many other states have been suggested in palaeo studies.

REPLY: This is clarified in the new version of the text.

L114: "Dynamic friction" needs explaining, or replaced with a simpler term (e.g. inertia, to match later discussion) - if too much detail for here, then leave to later

REPLY: This is clarified in the new version of the text.

L123: This is the first direct mention of socio-ecological systems – could introduce along with resilience theory above, given they are intertwined as concepts.

REPLY: A clarification/definition of social-ecological systems has been inserted, along with three new references.

"The social-ecological system (SES) framework (Berkes & Folke 1998) builds on the notion that nature is no longer simply the backdrop for social interactions, just as humans are not merely external forces acting upon ecosystems (Folke et al. 2011). Instead, social-ecological systems represent fully integrated, inter-dependent systems, where tightly linked feedbacks between social and ecological components

shape their overall behaviour and dynamics (Biggs et al. 2012).” The three references are new to the manuscript.

L128-2: Is this the same as the parameter described on lines 69-71? If so, I think it'd be more useful to define that here to make it easier for a reader unfamiliar with this sort of physics to follow, and if not then needs defining here (especially as forms an axis on later graphical results, so useful to explain at the outset to aid interpretation).

REPLY: This is a good point. We have strived to make it clearer. Furthermore, we have included a more representative Figure 2 to illustrate the existence of the metastable state to replace old Figure 2. (see also comment/reply above)

Eqn1: For clarity, best to redefine F as the free energy just before/after this equation.

REPLY: This is clarified in the new version of the text.

L143-146: This reasoning is good, but I think for readers not familiar with this sort of theory it might be worth unpacking here what this means / why this is useful (i.e. adding something the lines of "This allows us to project how the Earth system would behave in this model depending on it's initial state and subsequent trajectory with respect to the planetary boundaries"). The subsequent equations can then be clearly understood in the context of what's necessary to calculate this, even if the details are beyond the reader.

REPLY: This is implemented in the new version of the text.

L160: Clarify h_is.

REPLY: This is implemented in the new version of the text.

L165: It'd be clearer to state why it’s the case that adding these terms allows for metastable states, or leave mention of this it to the next section where you do discuss this.

REPLY: This is implemented in the new version of the text.

L175-177: This is interesting, but I think could be unpacked a bit more here if bringing it up.

REPLY: This is implemented in the new version of the text.

L194 / Fig2: This is implicit, but could explicitly describe here how this matches the Hothouse Earth hypothesis, in that it posited a Holocene attractor and a HE attractor, with some resilience to the latter that can be overcome with sufficient human perturbation (which also helps contextualise this result within that background literature).

REPLY: A new Figure 2 has been introduced to better illustrate the point.

L197-199: Would it be possible to explain how this is so in brief?

REPLY: This is implemented in the new version of the text.

L218: "Dynamic friction" here (and on line 114 & in abstract) presumably refers to the Rayleigh dissipation terms you just introduced, but this is not explicitly defined before using this term.

REPLY: This is clarified in the new version of the text.

L219: They are the properties that allow for ES resilience in this model, but they cannot (yet) be generalised to the real Earth system, given the assumptions made & limited observations.

REPLY: This point has been expanded in the new version of the text.

L220-229: This reads like a brief Discussion – might be better as part of Conclusion to end the paper on.

REPLY: This is implemented in the new version of the text.

L235-236: This is effectively the hypothesis of Steffen et al. (2018), but formalised in a thermodynamical model (whereas theirs remained conceptual) – would be good to contextualise your work here with that.

REPLY: This point has been expanded in the new version of the text.

L242-249: There may well be metastable states in the Earth system – as hypothesised by Steffen et al. (2018) for the future, or from palaeo/model studies for climate attractors in general, but I don't think it is currently possible to infer this from current observations (Steffen et al., for example, propose it as a hypothesis, but do not offer direct observations in support of it beyond some potential initial feedbacks), and no citation is given for this statement. You cite the latest PB assessment in the next paragraph, which might be the observations you're referring to, but as discussed above I think this measures distance from Holocene-like conditions, with that those conditions reflected being in a metastable state and that the bound of the state could be nearing being assumed in that framework. As such, while it is possible and worth investigating, I think it's more a case still of it being "unclear **if** the ES has a point of no return", and if so how close it is, and that modelling of this hypothesis should bear that in mind.

REPLY: This point has been expanded in the new version of the text.

L257: What does "internalising the workings of PBs" mean, in brief? Fine to direct reader to another paper for more details, but it's useful to catch the gist here without needing to.

REPLY: This point has been clarified in the new version of the text.

Again, we thank once again the reviewers for their constructive comments, which helped us to produce a better version of the manuscript.

With our best compliments,

The authors,

Orfeu Bertolami & Magnus Nyström