We would like to thank Maarten Ambaum (MA) for his critical, but constructive review. Below, we answer to all points raised.

[MA] This paper is written by a broad range of authors with a high level of expertise, who are well placed to provide an overview of thermodynamic optimality principles in earth system science. It is a well written and very readable paper of appropriate length, although I did feel that at points the paper suffers from a lack of detail, so that any reader who is not completely familiar with the background material does need to refer back to the primary literature on all topics. I personally enjoyed reading the paper, and enjoyed picking up on some of the topics that I was less familiar with.

Thank you for the kind words regarding readability of the manuscript. Regarding the lack of detail, we think it is always difficult to find that balance in a review paper. Reviewer 1, advised to be less comprehensive when describing the different studies. We will therefore critically look which literature needs to be described in more detail and which ones can do with less.

[MA] Reviews on this area have been written before, and I was looking at what this review adds above our present understanding; the paper promises quite a lot from the outset. I am afraid to say that in my opinion the paper does not deliver on these promises. In summary, the paper restates our lack of understanding about general properties of TOPs, but does not add any new insight.

Our aim was to find commonalities between successful applications in order to learn from them what is needed to apply TOPs. Subsequently, we aimed to clarify ambiguities around the term "maximum entropy production" and its use in many different contexts, sometimes leading to excessive room for interpretation and sometimes even to misleading results. By distinguishing between these different methods/forms we aimed to make clear that not all studies should be interpreted the same way. Based on input from Reviewer 1, we will focus in the revised version on choosing a few key examples to clarify concepts, but we also aim to demonstrate the generality of our arguments by putting all the relevant studies we could find into our concepts of explicit control volumes and energy and entropy budgets.

[MA] The first sentence does not at all reflect my understanding of the area: thermodynamic optimality principles, as addressed in the present paper, have little solid physical foundation. My reading of the literature, including some of my own contributions, is that many people attempt to demonstrate an underlying optimality principle in some given system; that is different from people using optimality principles to estimate model parameters or fluxes that are not already known or estimated from other physically explicit models. These optimality principles are invariably used post-hoc, and not as calculating principles. Of course, the authors admit as much in the final paragraphs of the introduction.

We guess that this comment refers to the first sentence of the abstract? We did not want to give the impression that TOPs have a solid physical foundation, since we agree with MA that they have not. The different ways TOPs are applied (demonstrating an underlying optimality principle vs estimating model parameters) are in our opinion supplementary to each other: The former one is needed to test if the principle can be used to get meaningful results, while the latter can only be done if one knows that the applied principle leads to meaningful results. We nevertheless agree with MA that in the literature often only one of the two ways is explored.

[MA] line 18 of the abstract we are promised that there is a correct and consistent use of the maximum power principle, which sounds farfetched to me. Maximum power is at best a hypothesis with a good amount of circumstantial evidence. However, it does not have the status of a physical theory which has a well-defined application area and procedure. At best we can hope to give a geography of cases where an application of maximum power appears to give a physically realistic result. Of course, any new results or understanding in the rest of the paper could prove me wrong, but I do not think such new results or understanding were provided; in fact the paper mostly is a descriptive geography of applications of TOPs. Fun to read, but probably not adding much insight.

With hindsight, we agree that this claim in the abstract is indeed too strong, and we will change this in the revised version where we will put more emphasis on synthesizing the results. Regarding the "geography of cases…", we want to emphasize that it was exactly our aim to investigate these cases to find commonalities between successful applications vs applications using different methods (but same terminology). As explained above, we will focus in the revised version on choosing a few key examples to clarify the different concepts.

[MA] Your third paragraph (p.2, l. 12-20) is a case in point: this is a great example of a falling object reaching terminal velocity in the presence of friction. It is used to point to the potential of thermodynamic optimality principles. Of course, the ultimate balance at terminal velocity is between production of heat at the expense of potential energy, and that seems to point to some possible thermodynamically optimal limit. But this does ignore the fundamental fact that, especially at higher Reynolds numbers, the primary, and limiting conversion is between potential and kinetic energy of the surrounding fluid. The balance is dynamic, not thermodynamic, evidenced by the fact that the energy conversion rate does not actually depend on the viscosity of the fluid, as long as the Reynolds number is large. The heat production is incidental; the terminal velocity is determined by a drag coefficient which is essentially a geometric property of the falling object.

It is indeed a good point that the surrounding fluid also gets kinetic energy. However, this kinetic energy is subsequently dissipated into heat as well. We will add this to the revised version. However, we do not agree that the "balance is dynamic, not thermodynamic", since this assumes that thermodynamics only deals with heat. Thermodynamics deals with all kind of energies (including mechanical energies) as is expressed in the first law of thermodynamics. Besides that this example of a falling object should be seen more as an analogue for than an example of application of a thermodynamic principle

[MA] Section 2.1: all basic and correct, but in section 2.2, and the rest I am surprised that there is never a reference to the Curzon-Ahlborn work on maximum power production which seems to me to address many issues of how heat is fluxed though a system using explicit models of conductive heat flux.

The Curzon-Ahlborn work on maximum power production is an extension of the classical reversible thermodynamics (on which the Carnot limit is based) to irreversible thermodynamics. But it still deals with heat engines where the heat reservoirs operate at fixed temperatures. Since our focus is on studies with flux-gradient feedbacks, the Curzon-Ahlborn work on maximum power production is out of the scope of this work. On top of that, none of the reviewed literature in section 3 and 4 referred to this type of work, so we don't feel that this should be part of this review.

[MA] The discussion around Eqs 3 and 4 imply a very particular physical set-up which is not at all explained in detail. For example: which flux is meant? The flux at the input terminal is typically different from that at the output terminal if mechanical work can be extracted. If the mechanical work is re-injected in the system, then we need to know at what temperature this happens. The schematic in Fig 1 implies this is re-injected at the output temperature (something that is not obvious at all: think about the thermodynamics around energy dissipation in tropical cyclones). I think you refer to this issue on p.21, l.15-21, but that was not particularly explicit. In other words, this part 2.2 does not explain much. Note also there are more explicit versions such as in Bejan's book and in my own book on Thermal Physics of the Atmosphere (Chapter 10.3) where you can find explicit expressions of "lost work" due to entropy production and of the effective temperatures of the input and output terminals. There you can also see that, for example, T_c depends on where the heat is lost from the system, so it is a geometric property of the system as much as a thermodynamic property. In that sense, T_c is not a function of J but of the whole of the fluid state. To take T_c as a function of J is a statement of belief, or approximation, but not a statement of a physical principle.

Thank you for pointing this out. Eq. 3 and 4 are connected to Fig. 1c-f, which we will explain better in the revised manuscript. The reason to start with the simple setups of Fig. 1 is to illustrate the general concept by choosing the simplest possible system expressing maxima in power and/or entropy production at certain conditions, i.e. a closed system consisting of two closed reservoirs at steady state, only exchanging heat and entropy between each other and with the surroundings. MA states correctly that T_c is a function of more than only J, but for the sake of simplicity, we stick with this relation (also because in the reviewed literature this assumption is generally made). We will add a sentence to state that T_c depends on more than J.

Please note that (as reviewer 1 pointed out) there were some flaws in Fig. 1, of which the major flaw is that work should be performed on the boundary of the system. Therefore, we will revise this figure: In the revised figure shown below (Fig. 1rev), we do include an explicit boundary around the two control volumes, clarifying that extracted power (G) refers to the work done on the boundaries and amounts to an extraction of energy (not entropy) from the system. We also clarify here that the heat produced by dissipation of the extracted work (D) has to enter across the same boundary across which power was extracted; hence it does not contribute to the internal entropy production. In this sense, Fig. 1 in the original manuscript was misleading. We will also accompany the revised version with an analysis of the energy and entropy balances of the general system in a Jupyter notebook, where all variables are explained with their units and where all necessary computations will be entirely transparent.

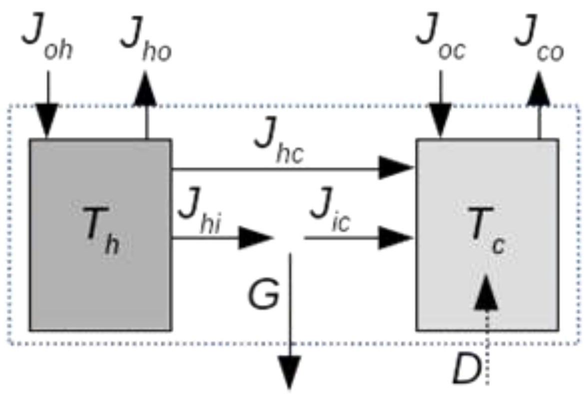

Figure 1rev: Revised 2-box model with J being heat fluxes [W], G is power [W], D is dissipation [W] and T is temperature [K]. The first letter of the subscript refers to the origin of the flux, while the second one refers to the destination of the flux: o is outside, h is hot, c is cold.

The most important message in section 2.2 is that we move from systems with fixed state boundaries (which are dealt with in classical textbook thermodynamics) to systems with fixed input rates. This is a crucial step to understand why (mathematical) optima exist, while also all literature reviewed in section 3 deals with this. We will stress this more clearly in the revised manuscript, while also explicitly referring to the differences with the classical textbook thermodynamics.

[MA] Just as an aside to p.9, l. 31, as part of a PhD project (thesis by J. Kamieniecki, University of Reading, 2019) we repeated the work by Herbert et al. (2013) and found that they do not offer the complete picture in their paper: the profiles become really rather unrealistic above the levels they plot in their paper. So the interpretation of this result is that the MEP principle did produce profiles that look somewhat realistic in a limited part of the vertical column, but as a whole are unrealistic.

Thank you for pointing this out. We will try to get hands on it to be able to refer to it.

[MA] You refer to our work on p. 12, l. 19-27; the description is correct. The ultimate failure of our work has to do with the fact that energy conversions were dominated by latent heat fluxes which scale more with the mean temperature of the system, rather than the temperature gradients. My suspicion is that we need to somehow exclude chemical conversions, such as described by Pauluis in several papers as the "Gibbs penalty" (such as in Kamieniecki et al., 2018, J. Atmos. Sci.).

Thank you for pointing this out. What we got out of the paper of Pascale et al. (2012a) was indeed that the material entropy production was dominated by the latent heat fluxes (i.e. chemical conversions), which do not feedback on the temperature gradient. So our explanation is in essence the same as MA is suggesting here. We will therefore stress in the revised version that latent heat fluxes are chemical conversions, which seem to be left out of the maximization.

[MA] The authors are honest in that they do not avoid the fundamental failure of TOPs, such as in Section 3.3, which in my view should be interpreted as: several applications give broadly sensible results, but they often do not survive deeper scrutiny, or broadening of application range. It looks like the initial set-up of the physical problem encodes the outcome, not the TOP itself.

We indeed wanted to be honest about it. However, we have to better include that in our synthesis.

[MA] In section 3.4 I must admit that I am not an expert on this literature, but I thought that non-linear chemical reactions have been widely used to explain pattern formation in nature, with an essentially thermodynamic argument: the free energy being a nonlinear function of some order parameter. Apologies if this sounds a bit vague, but I would have expected that a review, addressing pattern formation, would acknowledge that part of the physics literature which, as far as I understand it, is reasonably well established.

Thank you for pointing this out. We did not look into this part of the literature since our focus is on applications in Earth sciences. And since we will slightly change the focus in the revised manuscript to deriving the entropy balance for a set of cases we will not include this in the revised manuscript.

[MA] In section 3.5, I must again admit that this is not my specialist area of expertise, but the process described in p.17, l. 1 sounds very similar to the mechanism underlying the sandpile models of Per Bak, leading to SOC, which has a substantial body of literature around it. I may well be completely wrong here, but I would be surprised if there was no link between the SOC states and some appropriate TOP state in such models. At the end of Section 3.5, the authors, admirably, point again to an observed limitation of TOPs.

Thank you for the suggestion (MA probably refers to p16 (section 3.4)). A quick look told us that it has indeed some commonalities: especially the power law relations which are also present when describing some statistics regarding the derived Optimal Channel Networks (OCN). This will be made explicit in the revised version.

[MA] In Sec. 5.1 you discuss MP vs. MEP, and essentially argue they point in broadly the same direction. I agree with that, but it also means that, in the absence of a first principles physical theory for TOPs, we cannot decide at this stage which of the two options is the better one. It is clear that MEP has the potential to be related to MaxEnt, while MP might well be related to ideas around stationary action principles in Lagrangian or Hamiltonian descriptions of physical systems. There does not seem to be an overriding argument presented either way.

With hindsight, we agree that we could (and should) not have made this statement here.

[MA] Your discussion of minimum entropy production, as in Prigogine's work, seems to miss some clarity. I would have thought that minimum entropy production is a well established outcome for a system in the presence of linear fluxes (such as in Fourier's law) and fixed boundary conditions. The big transition to the kind of systems we are studying must then be the non-linearity of the fluxes, where flux values are linked to gradients in a non-trivial way, or the boundary conditions.

This is indeed correct. The example described here refers to an isolated system of two reservoirs with initially different temperatures. Since heat flows from the hot to the cold reservoir, the final state would be one in which the temperature of both reservoirs are equal and heat transfer has ceased (thermodynamic equilibrium). However, the minimum entropy production principle actually applies to much less trivial circumstances, involving coupled fluxes approaching a non-equilibrium steady state (see Kondepudi & Prigogine, 1998, Chapter 17.2). We will re-formulate the P22, L11-14 to:

"Besides these examples, systems close to thermodynamic equilibrium approaching steady state follow the principle of minimum entropy production (Prigogine, 1947). A trivial example of this is an isolated system of connected hot and cold reservoirs without energy transfer through its boundaries. Such a system evolves to a system where both reservoirs have equal temperature and the entropy production is zero (thermodynamic equilibrium). More interesting examples are given by e.g. Kondepudi & Prigogine (1998, Chapter 17.2)."

[MA] Your minimum requirements in section 5.3 appear mostly self-evident extractions from past experience, and therefore would not and should not exclude other ways that TOPs might be operational in the future.

Yes, we also realize that the way we described this is too firm. In fact it is an outcome of (or lessons learned form) the literature study. In the revised manuscript we will also formulate it as:

"Based on all these applications and our own simple analysis, we conclude that the MEP and MP principles are potentially useful if:

- the system is in a (quasi) steady state;
- there are sufficient flux-gradient feedbacks to result in non-linearity in power or entropy production;
- the system has enough degrees of freedom;
- the flux descriptions are physically based (related to thermodynamic forces);
- a complete balance of entropy and the exchanged quantities can be computed;
- the system definition results in non-trivial steady-state conditions.

If the above conditions are met, it is possible (under certain conditions) to find an optimal value of a flux coefficient that would lead to a maximum in entropy production by this flux or a maximum in extractable power from this flux. These optima are not necessarily the same and it is a matter of additional experimental and theoretical investigation to decide a priori if one of them is the correct attractor for a given system."

[MA] Your set of questions in section 5.4 address some of the fundamental issues, especially questions 2 and 4; these are issues that everybody working in the field has been aware of for many years; it looks like this review has mostly restated these issues, and by giving a list of applications with successes and failures has only restated the gaps in our understanding.

We are conscious that people well acquainted in this field are aware of these questions. Nevertheless, we think they are still useful to guide future research, while we also mention which studies are exemplary for that specific question, and which studies have already (partly) dealt with that question.

[MA] Your very final sentence then summarizes what I dislike about this review: "But this is only possible of the principle is applied in a transparent and correct way". This sentence implies there is such a thing as "the" principle, and that there is such a thing as the "correct" way. Both of these are highly disputed in the cited literature and the present review does not provide evidence for a solution for either of these.

We can see why you dislike it and, as said before, we agree that such a conclusion is too firm. Therefore we will state that the topic cannot finally be settled yet and that this is work in progress.

[MA] I found the paper well written, and have therefore not picked up on any typos. Here are three I did spot and did record: p. 2, l. 21: an –> a p. 20, l. 27: "a couple" means 2; you mention 6. p. 23, l. 13: fuzie –> fuzzy

Thank you for pointing this out. These will be corrected.

With kind regards,

On behalf of all coauthors,

Martijn Westhoff