We thank the reviewer for these helpful comments. Reviewer comments are listed below, along with our response to each. In some cases, we describe the proposed revisions to the manuscript (with line numbers) or refer to proposed revisions described in our responses to the other reviewers, but we recognize that the revised manuscript is requested in a subsequent step.

Comment 1:

The manuscript entitled "Reinterpreting the Budyko Framework" by Reaver and colleagues highlights several misconceptions regarding recent interpretations of results obtained using the Budyko framework. The authors especially criticize the common assumption that the wealth of functional Budyko curves represent expected trajectories through the Budyko space. By considering a stochastical model and observations from several hundred catchments, it is shown that catchment behavior in time does not follow the predicted trajectories within the Budyko space. The authors further highlight specific parameters used within parametric Budyko equations do not represent catchment-specific biophysical features. The authors thus conclude that Budyko-based results should be interpreted more carefully and thoughtfully.

Response 1:

We thank the reviewer for an accurate representation, though we note that another outcome of this work is to illustrate that the non-uniqueness of the parametric Budyko equations (i.e., there are several equally valid single parameter equations with different functional forms) fundamentally contradicts many recent interpretations of the parametric Budyko framework.

Comment 2:

The manuscript is generally in good shape, overall well structured and well written. The introduction motivates the study and provides an in-depth overview of the recent research within the field. It needs to be acknowledged that this paper addresses a somewhat heated debate on the interpretation and applicability of the Budyko framework in the context of biophysical drivers influencing the terrestrial water and energy balance. However, it is my assessment that the line of arguments and conclusions, as presented in this paper, are mostly adequate. The supporting data and examples seem valid, but I would appreciate a more in-depth justification of several assumptions.

Response 2:

We thank the reviewer for the kind words and the positive assessment of the manuscript. We address the request for additional justification of assumptions in subsequent responses.

Comment 3:

I have provided a relatively small number of comments. However, I would also like the authors to consider two more general remarks:

(i) I largely agree with the general conclusions of the paper. However, I know (from my own experience) that the debate on the interpretability of the Budyko framework is somewhat heated. Therefore, I think it needs to be noted that the Budyko framework remains a powerful concept when interpreted and applied correctly. And I don't think that you necessarily "reinterpret" the Budyko framework. I fully agree that it needs to be acknowledged that there has been a rather large number

of recent studies that overinterpreted results. Nonetheless, these studies still present results that are valid and sound within their specific setting. However, any interpretation going beyond these settings is not adequate, which needs to be acknowledged without condemning previous research. You already highlight this in your introduction, but I think you also need to be more careful throughout the rest of the manuscript (see some of the more specific comments below).

In this spirit, I would like to see a more positive evaluation of the Budyko framework per se. I think that the framework, given its adequate application and interpretation, remains super useful. It might thus help to better outline how upcoming Budyko-based research can profit from considering the limitations highlighted in this study. Your conclusions already provide some suggestions, but I still think that the Budyko framework has more potential besides being a global constraint (p. 25,l. 14), as it can also be applied within well-defined setups.

<u>Response 3:</u>

We agree that the Budyko framework is a powerful concept and have expressed its validity in the manuscript (page 25 lines 8-9, page 25 lines 14-16). However, we take the reviewer's point and therefore propose specific edits (listed at the end of the response) to strengthen this concept in the manuscript as well as outline useful directions for future Budyko-based research (see also proposed edits in the response to **Reviewer 2 Comment 25**). In addition, we do not intend to condemn previous research and explicitly attempt to convey this sentiment in the manuscript (page 3 lines 15-17) (see also our responses to **Reviewer 2 Comment 10** and **SC1 Comment 2**). However, we acknowledge the need to be more specific and state that we believe that only the "original" Budyko framework (i.e., the observation that the aggregate behavior of multiple catchments consistently produce a distinctive pattern in Budyko space) remains intrinsically useful. Specifically, the emergent aggregate Budyko curve pattern provides: (1) an empirical constraint for process-based model validation and theoretical investigations and into catchment hydraulic processing; (2) a practical constraint for process-based model calibration; and (3) allows for probabilistic predictions of $\bar{E}$ and $\bar{Q}$ as well as changes in $\bar{E}$ and $\bar{Q}$ for ungauged basins with limited data. Used in these contexts, the Budyko framework is powerful and useful.

However, based on our review of the literature (and the analyses and arguments presented in this manuscript), we do not agree that the parametric Budyko framework is intrinsically useful to the catchment hydrology community. Since the functional forms of the parametric Budyko equations are not empirically valid (see Sections 3.1.2 and 3.2.2 in the main manuscript), these equations are literally only an arbitrary mathematical transform of the evaporative index. As we state in the manuscript (page 21 lines 18-19), with this mathematically sound interpretation, use of the parametric framework in hydrological applications is theoretically benign, but it is unnecessary and can easily amplify errors. Essentially all studies that use single-parameter Budyko equations could have conducted the same analyses on $\bar{E}$ or the evaporative index instead.

A good example of a study that used the parametric Budyko framework robustly, appropriately, and in a manner which did not impact its outcome is Greve et al. (2020). This study calibrated several parameters of a global hydrological model using a parametric Budyko framework-based constraint. This methodology significantly improved the global hydrological model's performance compared to its uncalibrated version. However, as we argue in our manuscript (see page 21 lines 7-17), Greve et al.

(2020) could have compared the simulated and empirical joint distributions of $\frac{\bar{E}}{\bar{P}}$ and $\frac{\overline{E_0}}{\bar{P}}$ directly, without using the distributions of the catchment-specific parameter (ω) as an intermediary. The parametric Budyko framework could be easily removed from this analysis without changing the results. Thus, while the use of parametric Budyko framework did not change the outcome of the study, it acted as an unnecessary mathematical transform. Similar benign situations occur for essentially all appropriate uses of the parametric Budyko framework, while inappropriate uses often lead to spurious results.

Despite this critique, we also stress that, even for prior research that utilized inappropriate interpretations of the parametric framework, both the intent and much of the effort of previous research can be preserved. For example, as we state in the manuscript (page 21 lines 9-14), any study that has related n or w to catchment biophysical features could remove the parametric framework from their analyses and use their same analytical tools to relate $\bar{E}$ or $\frac{\bar{E}}{\bar{P}}$ to biophysical features directly. This would preserve most of the analyses of such studies (i.e., same analytical methods) as well as the intent (i.e., understanding the interactions between $\bar{E}$ and catchment biophysical features).

To strengthen our representation of the validity of the "original" Budyko framework in the manuscript, highlight how the intent and effort of previous research can be preserved, and to better outline useful directions for future Budyko-based research, we propose the following edits:

1) Add the following sentence to page 3 line 17:
"Additionally, we emphasize that the Budyko framework based on the curve-like clustering pattern observed across multiple catchments is a powerful and useful concept when used appropriately and with proper context."

2) Edit the sentence page 25 lines 14-16 to the following:
"Additionally, to be a valid representation of catchment evapotranspiration, process-based models need to able to reproduce the empirically established general Budyko curve behavior (i.e., nonparametric) when applied to multiple catchments across a range of climates. As such, the general Budyko curve behavior can serve as a global constraint (i.e., calibration or validation) in the application of such models, e.g., Greve et al. (2020). Furthermore, while the parametric Budyko framework lacks predictive power, the nonparametric framework allows for probabilistic predictions of $\bar{E}$ and $\bar{Q}$ as well as changes in $\bar{E}$ and $\bar{Q}$ for ungauged basins. Within these contexts, the nonparametric Budyko framework is a useful conceptualization. "

3) Add the following to the end of Section 4.2.1:
"While the acknowledgment of the proxy nature of the catchment-specific parameter and $\frac{\bar{E}}{\bar{P}}$ casts doubt on the specific conclusions of previous parametric Budyko-based research, we note that both the intent and much of the effort of many such studies can be preserved. For example, studies that related n or w to catchment biophysical features using various analytical tools could employ the same methods relate $\bar{E}$ or $\frac{\bar{E}}{\bar{P}}$ to biophysical features directly. Doing so would preserve most of the analyses of such studies (i.e., near identical methods) as well as their intent (i.e., understanding the relationship between $\bar{E}$ and catchment biophysical features)."

Comment 4:

(ii) Your theoretical example using Porporatos model is neat. However, it is still an artificial example and also needs to be interpreted as such. You use one model (Porporatos model) to investigate the characteristics of another model (Budyko). Fine, but you need to thoroughly justify that Porporatos model is an appropriate choice in this context: Is the choice of parameter values for the different cases realistic? What kind of conditions do these parameter values represent? Is there any real-world example that would illustrate your choice?

Additionally, as the parameters might be independent within your theoretical modeling framework, they might not be independent under real-world conditions. That also represents another problem of the large number of studies trying to identify biophysical controls. There is no single parameter that controls the partitioning of precipitation into evaporation and runoff. It is rather a convoluted mess of different processes that interact with each other.

Response 4:

We agree that the Porporato model is an artificial example (i.e., model investigating a model) and has some limitations (e.g., assumes E_0 is constant), though we do not believe that these properties influence the general conclusions from the theoretical test. Our primary justifications for using the Porporato model were its simplicity (see page 10 lines 20-25 and page 11 line 1), physically-based nature (see page 10 lines 7-9), and its previous use within the Budyko framework literature (see page 10 lines 5-7). However, we take the reviewer's point that the appropriateness of the Porporato model's use could be further justified in the manuscript and propose to do so through the suggested edits in this response.

We understand the reviewer's concerns regarding the model's parameter values and whether they are reflective of conditions in real catchments, as models should typically reflect reality. However, for the Porporato model, this is somewhat unimportant since all parameters appear in ratios. This means the individual parameter value magnitudes are less important than the relative magnitudes between parameters. This flexibility allowed us to choose values that were more useful for illustrative purposes as opposed to precise values reflective of a particular catchment conditions. Specifically, the values of the parameters were primarily chosen for the following reasons: (1) to maintain the simplicity of the illustrative examples (e.g., using integers); (2) to allow all of the test cases to be expressed through a single functional relationship (see page 11 lines 12-15); (3) to produce trajectories that would be visually informative (e.g., not restricted to a small portion of Budyko space such as being compressed at the water and energy limits). However, we note and emphasize that choosing values representing particular catchment conditions (e.g., an effective rooting depth of 0.374 m) would not change the results (since the same trajectories could be produced by adjusting other parameters), but the simplicity and clarity of the test cases would be lost. To reflect this view in the manuscript better, we propose to add the following paragraph to the end of Section 3.1.1:

"The effective climate and landscape parameters exclusively appear in ratios within the Porporato model. This means only the relative magnitude between parameters is important. Therefore, for our test cases, we chose parameter values to maintain illustrative simplicity, allow all test case to be expressed in a unified functional form, and to produce visually informative trajectories not restricted to a small portion of Budyko space. This choice does not impact the conclusions of the test cases, since the ratio

nature of the model's parameters means the exact same trajectories can be made from infinitely many different parameter combinations."

The reviewer rightly points out that there may be possible dependencies between model parameters and that the test cases should correspond to real-world examples. We note that for the variable parameter test cases, we explicitly considered the possible dependencies of the climate parameters (i.e., the relationships between η and ψ) (see page 11 lines 12-15) and chose test cases that are reflective of real-world conditions (see page 11 lines 7-12 and the references included, Trenberth (2011) and Fischer et al. (2014)).

Comment 5:

Minor comments:

p. 3, l. 7: Please also consider Padron et al., 2017. It provides a comprehensive overview of inconclusive and contradictory evidence obtained from using eq. 6.

Response 5:

We agree that Padrón et al. (2017) should be included and propose to add the following sentences starting at page 3 line 5:

"Furthermore, Padrón et al. (2017) undertook a comprehensive overview of the wide variety of biophysical features proposed to control the catchment-specific parameter, finding that most proposed features did not actually correlate with the parameter and the types of features that were correlated varied sustainably between climatic regions."

Comment 6:

p. 4, l. 4-7: It will be helpful to already mention those equations here.

Response 6:

We agree that it would be helpful to have already introduced both of the parametric equations for readers who are uninitiated to the catchment hydrology literature. However, doing so would significantly alter the flow and structure of the introduction and background, and such readers are only a portion of the intended audience. Given the apparent contentious status of the topics covered into this manuscript (as the reviewer notes in **Comment** 2 and **Comment 3**), we have intentionally structured our introduction to ensure that the reader is informed of the full context, implications, and summary of the manuscript's contents before being introduced to the more detailed background, analyses, and discussions. We propose to keep this structure, however we understand the reviewer's point and therefore propose to edit the sentence on page 4 lines 4-5 to:

"We argue and demonstrate herein that the two widely accepted parametric Budyko equations (i.e., those derived in Zhang et al. (2004) and Yang et al. (2008)) are non-unique, meaning they are only two of many possible single-parameter Budyko equations."

Comment 7:

Sec. 2.3: I think it might help to incorporate this section into Sec. 2.1?

Response 7:

We understand that “Budyko’s interpretation of explicit curves” could naturally be classified as part of an “Overview of the Budyko hypothesis and equations”, however, we specifically placed it after “Current interpretations of explicit Budyko curves and the parametric framework” to contrast how current interpretations have evolved (or strayed) from the “original” intent of explicit curves. This section provides the start of our reinterpretation (or “retrospective interpretation”) from the current state of the framework. Therefore, we believe it is important to highlight Budko’s interpretation independently and following the introduction of current interpretations.

Comment 8:

p. 10, l. 14: It might be helpful to further explain what you mean by "Budyko-like"?

Response 8:

We propose changing p. 10, l. 13 to:

“We first write the model of Porporato et al. (2004) in a form which can be plotted in Budyko space,…”

Comment 9:

p. 10, eq. 8: Could you explain in more detail how you estimate the aridity index from this equation?

Response 9:

We propose adding the following sentence at Line 1 on Page 11 to make this clearer:

“The ratio of ψ and η is the aridity index, $\phi = \frac{\overline{E_0}}{\bar{P}} = \frac{\overline{E_0}}{\alpha\lambda} = \frac{\psi}{\eta}$.”

Comment 10:

p. 11, l. 15: Why 2m? Well, this is related to my second major comment above. The choice of these parameter values needs to be justified. What kind of soil characteristics does Z0=2m represent? I know that your overall conclusions won’t change when setting Z0=1.9, but it is important to understand what it means and what kind of real-world characteristics your choice represents.

Response 10:

We have addressed this comment in our response to **Comment 4.**

Comment 11:

Sec. 3.2: Why don’t you include this subsection in the Background part (Sec. 2.)?

Response 11:

The Background section currently only contains existing information in the literature. We believe the content of Section 3.2 is a new contribution and therefore the content belongs in the methodological and discussion portions of the manuscript.

Comment 12:

p. 14, l. 8-10: Is that true? Are Eqs. 5 and 6 considered the only valid parametric Budyko equations? Do you have more evidence for this statement?

Response 12:

We thank the reviewer for addressing this point. Upon review of our original claim, which appears to be a common theme in the literature, we now agree that there are other single parameter functional forms that also satisfy the uniqueness requirement, e.g., a form of the equation introduced in Porporato et al. (2004). As such, we have proposed edits to Section 3.2.2 in our response to **Reviewer 2 Comment 20** to focus more on the properties that are typically used to justify the validity of Eq. (5) and (6) and highlight how these properties are not unique to them.

Comment 13:

Figure 2: I know it is hard to convey all the necessary information into one Figure, but I have to admit that this one is especially difficult to interpret. The trajectories are a big mess (and to a certain extent this is exactly what you want to highlight here). However, Figure 3a is of more value in this context. If you like to keep Figure 2, maybe consider drawing thinner red lines or introduce some transparency?

Response 13:

We agree with the suggestions of the reviewer and those provided in **Reviewer 2 Comment 23**. We propose to reduce the thickness of the red lines and logarithmically (base 10) scale the abscissa and modify the text referencing this figure accordingly. The resulting figure is:

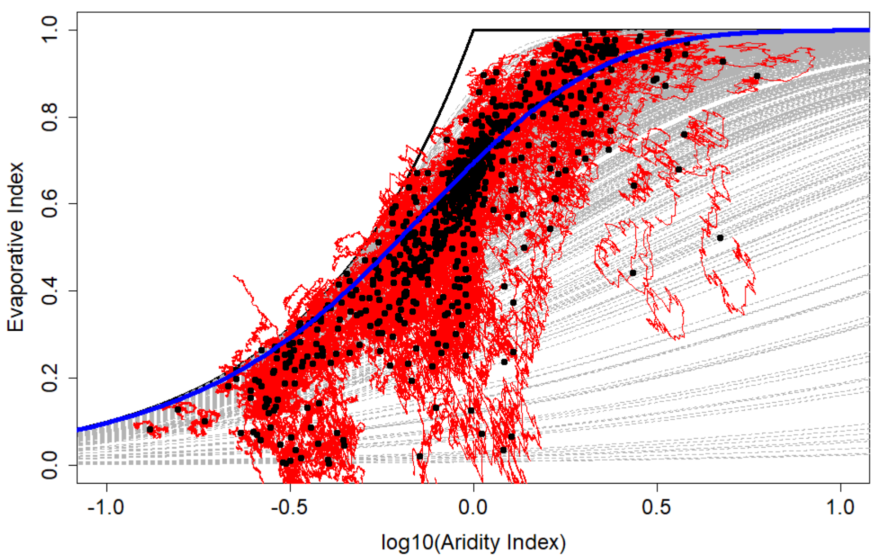

Figure 2: Semi-log plot of the Budyko space locations (black dots) of the 728 UK and US reference catchments and their corresponding expected Budyko curve trajectories, Eq. (5) (gray dashed curves) and 10-year average actual trajectory realizations (red solid curves). The global behavior of the catchments and their actual trajectories generally agrees with the the non-parametric Budyko, Eq. (3) (blue solid curve) but not the expected parametric Budyko curve trajectories.

Comment 14:

p. 25, l. 1-2: Maybe this would be a better title: A reinterpretation of explicit Budyko curves and parametric Budyko equations.

Response 14:

We agree this title provides a more focused representation of the contents of the manuscript. Therefore, we propose adopting a modified version of it in the revised manuscript:

"A reinterpretation of explicit curves and parameters of the Budyko framework"

Comment 15:

p. 25, l. 20-26: I agree that the interpretation of the parameter representing landscape features is misleading. Calling it a catchment-specific parameter is not justified either. However, even though it is a lumped parameter just existing as a mathematical necessity without any a priori physical interpretation, there might still be an a posteriori physical interpretation. You call the parameter a proxy variable for E/P, which is, in fact, also some sort of physical interpretation. That means, if you assume a constant aridity index and change E/P, the parameter changes as well. Vice versa, if you change the parameter, E/P changes as well. I think the misleading interpretation here is often more related to the assumption that the parameter somehow controls E/P, which is definitely not true.

Response 15:

As we state in the manuscript (page 13 lines 15-29), in order for the "catchment-specific parameter" to have *a posteriori* physical interpretation, the associated functional form (i.e., Eqs. (5) or (6)) would have to be empirically valid. But, because our empirical test of the catchment trajectory conjecture refutes the parametric Budyko curves' empirical validity, making Eqs. (5) and (6) under-determined (i.e., 1 equation with 2 unknowns, $\frac{\bar{E}}{\bar{P}}$ and n or w), we conclude (page 20 lines 16-29) that the "catchment-specific parameter" does not have an interpretation independent of $\frac{\overline{E_0}}{\bar{P}}$ and $\frac{\bar{E}}{\bar{P}}$.

We say the "catchment-specific parameter" is a proxy for $\frac{\bar{E}}{\bar{P}}$ since, in practice, the values of $\frac{\overline{E_0}}{\bar{P}}$ (usually taken as a known quantity) and $\frac{\bar{E}}{\bar{P}}$ ***always*** determine the "catchment-specific parameter". For real catchments, the association between $\frac{\bar{E}}{\bar{P}}$ and the "catchment-specific parameter" is one-way; $\frac{\overline{E_0}}{\bar{P}}$ and $\frac{\bar{E}}{\bar{P}}$ are used to compute n or w. It is never the case that $\frac{\overline{E_0}}{\bar{P}}$ and n or w are used to compute $\frac{\bar{E}}{\bar{P}}$. Therefore, we agree with the reviewer that the "catchment-specific parameter" does not control $\frac{\bar{E}}{\bar{P}}$, however $\frac{\bar{E}}{\bar{P}}$ completely controls the value of the "catchment-specific parameter" (page 21 lines 1-7).

Comment 16:

p. 25, l. 27-28: This statement is too strong in my opinion (see also my first major comment). Any interpretation of obtained results is valid within their specific setting. However, it is the overinterpretation and generalization that is "untenable" (which is a very strong word in this context).

Response 16:

We have largely addressed the philosophical underpinnings of this comment in our response to **Comment 3**. While we believe that the current sentence is accurate, we take the reviewer's point about the strong wording of this statement. Therefore, we propose to edit the sentences on page 25 lines 27-29 to:

"The collective results from our analyses suggest that current commonly held interpretations of Budyko curve trajectories and the parametric Budyko equations are unsupported. We propose that the catchment hydrology community look critically at the well-accepted but unjustified interpretations that are the current commonly held standard."

References:

Fischer, E. M., Sedláček, J., Hawkins, E., and Knutti, R.: Models agree on forced response pattern of precipitation and temperature extremes, Geophysical Research Letters, 41, 8554-8562, 10.1002/2014gl062018, 2014.

Greve, P., Burek, P., and Wada, Y.: Using the Budyko Framework for Calibrating a Global Hydrological Model, Water Resources Research, 56, e2019WR026280, 10.1029/2019wr026280, 2020.

Padrón, R. S., Gudmundsson, L., Greve, P., and Seneviratne, S. I.: Large-Scale Controls of the Surface Water Balance Over Land: Insights From a Systematic Review and Meta-Analysis, Water Resources Research, 53, 9659-9678, 10.1002/2017wr021215, 2017.

Porporato, A., Daly, E., and Rodriguez-Iturbe, I.: Soil water balance and ecosystem response to climate change, The American Naturalist, 164, 625-632, 2004.

Trenberth, K. E.: Changes in precipitation with climate change, Climate Research, 47, 123-138, 10.3354/cr00953, 2011.

Yang, H., Yang, D., Lei, Z., and Sun, F.: New analytical derivation of the mean annual water-energy balance equation, Water Resources Research, 44, n/a-n/a, 10.1029/2007wr006135, 2008.

Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H. S., Western, A. W., and Briggs, P. R.: A rational function approach for estimating mean annual evapotranspiration, Water Resources Research, 40, n/a-n/a, 10.1029/2003wr002710, 2004.