We have provided point by point responses below with author replies in bold.

The authors have used a physically based hydrological model to improve waterbudgeting at catchment scale. In particular, they have considered the original Budyko model as a reference and shown that by accounting for ground water inflow/outflow, water budgeting can be done more accurately. I really appreciate the authors' effort to undertake such an extensive numerical analysis. The article looks suitable for publication although I think a couple of key concerns the authors need to address.
We thank the referee for their support of our work and appreciate their suggestions. We have provided detailed responses to each point below.

1. Purpose of the study
The authors need to elaborate on the usefulness of their study. The physically based hydrological model they are using has many parameters; they cannot take that model to a random ungauged catchment and predict its hydrological variables. On the other hand, the Budyko model is a universal deterministic model which can be applied to any ungauged catchment. It is thus not surprising that the multi-parameter model will perform better after calibration. I don't think their study is very informative unless they integrate a deterministic physically based hydrological model with the Budyko model to improve prediction.
We agree with the limitations noted by the reviewer and would like to clarify that the purpose of our study is not to predict shape parameters or evapotranspiration in ungauged basins. Rather we are using the model as a controlled numerical experiment to demonstrated they ways that storage changes will influence Budyko relationships across a broad range of physical settings and for various experimental approaches. The purpose here is to better inform other studies that seek to attribute variance in the Budyko space to physical watershed characteristics but that may be lacking the required data to account for groundwater surface water interactions. We have revised the manuscript significantly to better focus on these goals throughout and especially in the abstract introduction and conclusions. For example, we have added the following text to the end of the introduction to clarify this point:

> ***"…Groundwater observations sufficient to precisely characterize watershed storage changes are difficult to obtain and are not widely available. Therefore, it seems unlikely that groundwater storage calculations will be added into most Budyko analyses; more work is needed to understand the sensitivity of Budyko relationships to changes in storage and the implications of assuming of no storage changes without the ability to regularly verify this assumption.***
>
> ***We have identified three main approaches to estimate evapotranspiration (E) in Budyko analysis: First, if it's not possible to measure E directly, it is commonly estimated as the difference between precipitation and river outflow in a basin. Second, some studies measure E directly using a variety of field methods. Third, as is the case with the more recent studies that seek to account for storage changes, observed E values are augmented with measurements of groundwater surface water exchanges to estimate the 'effective precipitation' that is available for surface processes (i.e. outflow and E).***
>
> ***Here we hypothesize that storage changes will bias Budyko results in predictable ways, as has been indicated by previous studies, but that the direction of the bias will vary based on the way that evapotranspiration is handled within a study. We evaluate this hypothesis by comparing Budyko relationships generated following the three different approaches using the***

outputs of a physically based hydrologic model that directly simulates the integrated groundwater surface water system over a large spatial domain at high resolution. The three primary goals of our comparative analysis are as follows:

1. ***Evaluate the sensitivity of Budyko relationships to groundwater storage changes***
2. ***Characterize systematic differences in the impact of storage changes on Budyko relationships***
3. ***Illustrate variability between approaches across physical settings and spatial scales"***

2. Clarity of presentation.

It is quite hard to follow what the authors are saying at many places. In my opinion, the presentation needs to be simple. If the authors' objective is to show how the physically based hydrological model is doing a better job at water budgeting, they need to focus on that part more. There is not a single figure showing a direct comparison between prediction by the physically based hydrological model and that by the Budyko model... Terms need to be defined prior to their usage. For example, in Line 27 the authors are talking about Budyko curve parameters. The authors are actually talking about Fu model's parameters (Budyko model does not have any parameter).

We have significantly revised the manuscript following this comment, as well as the comments from other referees, to try to improve clarity. We have revised the abstract and this text is no longer included. As noted above we are now much more explicit about our goals in the introduction. Also, we reorganized the methods section to include all of the details of the three approaches, expanded the discussion of the results figures to be more descriptive and we shortened the conclusions section to remove redundant material and focus back on the original goals of the paper.