Response to Reviewers' comments

We greatly appreciate the reviewers providing valuable and constructive comments on our manuscript HESS-2020-548. We seriously considered each comment and will revise/improve the manuscript accordingly. The individual comments are replied below. In the following the reviewer comments are black font and our responses are blue and to assist with navigation we use codes, such as R1C2 (Reviewer 1 Comment 2).

Anonymous Referee #1

R1C1: This manuscript studies the effect of elevated atmospheric CO2 on runoff at the catchment scale. The approach is based on a combination of models linking elevated CO2 to plant water demand (mediated by leaf area and stomatal conductance changes) and supply (depending on soil water access via changes in rooting depth). The approach is to my knowledge novel (despite building on several previous models and data analyses) and results are interesting. The topic is certainly suitable for HESS. However, I have some concerns regarding the theoretical setup of this work, specifically how different models have been linked and the consistency of underlying modelling assumptions.

Reply: Thanks for your encouraging and constructive comments. Your individual comments are replied to below.

R1C2: Consistency across stochastic soil water balance models. The model by Guswa (2008) assumes that actual evapotranspiration (ET) is fixed and equal to potential ET (PET) as soil moisture varies between the wilting point and saturation. In contrast, the model by Porporato et al. (2004) assumes that actual ET increases from 0 at the wilting point to PET at saturation. These two models are therefore based on different assumptions regarding the relation between actual ET and soil moisture, which in turn affect the long-term mean soil moisture and actual ET values. As a result, the ET/precipitation vs. PET/precipitation relations (i.e., relations in the Budyko space) will differ between these models. To develop a self-consistent theoretical approach to study elevated CO2 effects on runoff, a single stochastic soil moisture model should be selected and used throughout. For example, see how the model by Porporato et al. (2004) can be integrated into Guswa's framework for rooting depth (Guswa, 2010, doi:10.1029/2010WR009122).

Reply: We agree with this reviewer on the raised issue. In fact, we realized it when building the BCP model in 2012. The reason that we still go for Guswa-2008, instead of Guswa-2010, is that the solution of transpiration (T) in Porporato-2004 includes an incomplete gamma function with rooting depth contained in both parameters of that incomplete gamma function. This feature makes the analytical solution of dT/dZ_r extremely complex (see below equation) and it is almost impossible to derive an explicit solution for Z_r. We believe this is the reason that Guswa did not provide an explicit solution of Z_r in his 2010 paper. The results presented in Guswa-2010 and

Porporato-2014 were derived numerically but only for specific cases (e.g., with specified aridity index or T_P or dT/dZ_r).

$$
\begin{aligned}
&dT/d\gamma = -T_P * ((exp(-\gamma)*(\gamma^{(W*Zr-2)}*(W*\gamma-1)+W*\gamma^{(W*\gamma-1)}*\ln(\gamma)))/(\Gamma(W*\gamma)-\Gamma(W*\gamma,\gamma)) \\
&-(\gamma^{(W*\gamma-1)}*exp(-\gamma))/(\Gamma(W*\gamma)-\Gamma(W*\gamma,\gamma)) \\
&+(\gamma^{(W*\gamma-1)}*exp(-\gamma)*(expint(1-W*\gamma,\gamma)*(W*\gamma*\gamma^{(W*\gamma-1)}+W*\gamma^{(W*\gamma)}*\ln(\gamma)) \\
&+\gamma^{(W*\gamma)}*(W*(hypergeom([W*\gamma,W*\gamma],[W*\gamma+1,W*\gamma+1],-\gamma)/(W^2*\gamma^2) \\
&+(pi*(log(\gamma)-psi(1-W*\gamma)+pi*cot(pi*(W*\gamma-1))))/(\gamma^{(W*\gamma)}*sin(pi*(W*\gamma-1))*\Gamma(1-W*\gamma))) \\
&-expint(-W*\gamma,\gamma))-W*\Gamma(W*\gamma)*psi(W*\gamma)))/(\Gamma(W*\gamma)-\Gamma(W*\gamma,\gamma))^2)
\end{aligned}
$$

and $$\gamma = \frac{Z_r * \mathrm{SWHC}}{\beta}$$

where SWHC is soil water holding capacity and β is the mean rainfall intensity.

In the BCP model, the Guawa's model is used to estimate the effective rooting depth, which is then used to calculate the Porporato's parameter ω (the symbol γ is used in Porporato-2004). According to Guswa-2010, the Porporato's solution for transpiration will lead to a slightly deeper rooting depth than the Milly's solution for transpiration (adopted in Guswa-2008 and this study). Despite that, the responses of Z_r to changes in climate are essentially the same when the two transpiration solutions are adopted. Moreover, the responses of Z_r to changes in CO_2 in the two solutions should also be essentially the same, since the effects of CO2 on Z_r are expressed via water use efficiency and potential transpiration in our parameterization, which are independent of Z_r parameterizations. In summary, using different transpiration solutions (Milly-1993 versus Porporato-2004) would only lead to difference in the resultant absolute magnitude of runoff (Q) but unlikely to result in differences in the response of Q to CO_2 changes in any notable way, especially when the relative magnitude is used.

We will discuss this point in the revised manuscript.

R1C3: Budyko curve parameterization. The authors use results from Porporato et al. (2004) to link the exponent n in Eq. (1) to rooting depth, water holding capacity, and mean precipitation event depth. This approach is based on analysis of "data from Porporato et al. (2004)" (L103), though it is important to emphasize that in that paper there are no data (except for net primary productivity), so the regression reported in Eq. (2) is obtained by fitting results from the analytical model in Porporato et al. (2004). This step is quite unnecessary, since the results are already in a close-form solution, which can be used directly without any fitting. In other words, Porporato et al. (2004) already provides a fully parameterized Budyko curve, which should be used for consistency with the other parts of the model instead of Eq. (1).
Reply: We agree with this reviewer from a theoretical perspective. However, from a practical perspective, Porporato's model is much more complex than the Budyko model and can only be solved numerically (for the reason stated in the reply to R1C2).

With a specified model parameter, Porporato et al. (2004) proved the similarity between their solution and the Budyko's solution of mean annual water balance. Compared with Porporato et al (2004), the Budyko's formulation is much simpler, which allows an analytical attribution of Q changes. Therefore, developing relationship between the Budyko's parameter (here the Choudhary's expression of the Budyko curve) and Porporato's parameter is a simple yet effective way to solve the problem. The same approach has been adopted in previous studies (e.g., Donohue et al., 2012, https://doi.org/10.1016/j.jhydrol.2012.02.033; Liu et al., 2016, https://doi.org/10.1016/j.jhydrol.2016.10.035; Yang et al., 2016, https://doi.org/10.1002/2016WR019392; Shen et al., 2017, https://doi.org/10.1016/j.jhydrol.2017.09.023; Zhang et al., 2018, https://doi.org/10.1002/2017WR022028).

This reviewer was correct that there were no data in Porporato et al. (2004). What we obtained from the authors of that paper (via email exchange in 2010) is their numerical solutions of the corresponding E/P for every 0.1 increment in PET/P for six separate γ curves. By numerically solving the Choudhary's formulation of the Budyko curve, we determined the values of the Budyko parameter (n) that correspond to the E/P values of each of the six γ curves. We then pooled all $n-\gamma$ pairs together to derive a simple relationship between them (Eq. 2 in the manuscript).

R1C4: Model interpretation at annual time scale. The models by both Porporato et al. (2004) and Guswa (2008) have been developed for growing season conditions, assuming no seasonality in precipitation and potential evapotranspiration. In this contribution, these models are interpreted as representative of the whole hydrologic year and used to partition variability in annual runoff. I wonder if and how the original model assumptions and the current model interpretation can be reconciled.
Reply: In our study, the Guswa's model was indeed applied for growing season to determine the effective rooting depth (Line xxx). The determined effective rooting depth during growing season is then used to determine the Porporato's parameter and further, the Budyko parameter. It should be noted that the effective rooting depth is essentially the maximum depth of hydrologically active soil layer, which should remain unchanged between the growing season and the whole hydrologic year.

R1C5: Role of precipitation event frequency. Eq. (2) neglects the effect of precipitation event frequency on the shape of the Budyko curves from Porporato et al. (2004) framework. The variations in frequency across climates can be more pronounced than variations in mean event depth.
Reply: We agree with this reviewer that the precipitation event frequency is important in the control of precipitation partitioning into evapotranspiration and runoff and the Porporato et al (2004)'s framework does not explicitly account for it in their model parameter. Nevertheless, the Porporato's framework considers both the total precipitation amount the mean event depth, which when combined provide information about event frequency. Therefore, the effect of variation in event

frequency on the hydrological partitioning in the Porporato's framework and the BCP model is implicitly expressed by the variations in both total precipitation amount and mean event depth.

R1C6: Interpretation of results from Donohue et al. (2013). Eq. (6) presents an iterative scheme to estimate changes in WUE through time, but in the original articles by Donohue et al. (2013, 2017) steady state models are developed, without an explicit dynamic component. The time scales to achieve steady state are probably in the order of decades (necessary for vegetation change), not years as indicated in Eq. (6).
Reply: Eq. (6) follows the gas-exchange theory at the leaf-level to quantify the response of WUE (W in the following equations for simplicity) to elevated CO_2, originally given by Wong et al. (1979),

$$W_L = \frac{A_L}{E_{TL}} = \frac{g_s(C_a - C_i)}{1.6 g_s (v_i - v_a)} = \frac{C_a}{1.6v}\left(1 - \frac{C_i}{C_a}\right)$$

where A (g C m^{-2} s^{-1}) and E_T (mm s^{-1}) stand for the assimilation and transpiration rate, respectively, and the subscript L denotes the leaf-level variables. C_a (ppm) and C_i (ppm) respectively represent the ambient and intercellular concentration of CO_2, and v_a (Pa) and v_i similarly represent ambient and intercellular concentration of water vapor while g_s (m s^{-1}) is the stomatal conductance to CO_2. The numeric factor 1.6 accounts for the greater diffusivity of water vapor relative to CO_2 in air [Wong et al., 1979]. We use v to denote the leaf-to-air water vapor pressure difference (Pa), which is approximated by the atmospheric vapor pressure deficit in subsequent analysis. The relative change in W_L is given by:

$$\frac{dW_L}{W_L} = \frac{dA_L}{A_L} - \frac{dE_{TL}}{E_{TL}} = \frac{dC_a}{C_a} - \frac{dv}{v} + \frac{d\left(1 - \frac{C_i}{C_a}\right)}{\left(1 - \frac{C_i}{C_a}\right)} .$$

Observations have shown that for a given photosynthetic pathway (*i.e.*, C3 or C4 species), C_i/C_a is relatively conservative [Arens et al., 2000, Long et al., 2004, Wong et al., 1979]. The response of the term $1 - C_i/C_a$ to a change in v can be quantified by taking $1 - C_i/C_a$ as being approximately proportional to the square root of v [Donohue et al., 2013; Farquhar et al., 1993, Medlyn et al., 2011]. Therefore, Eq. (2) can be written as:

$$\frac{dW_L}{W_L} = \frac{dA_L}{A_L} - \frac{dE_{TL}}{E_{TL}} \approx \frac{dC_a}{C_a} - \frac{1}{2}\frac{dv}{v} .$$

Eq. (6) in the current manuscript is essentially the same as the above equation. This equation does not require steady-state to be satisfied. However, the above equation is for leaf-level fluxes. Applying this equation at the canopy-scale implicitly assumes the same upscaling factor when converting the leaf-level assimilation and transpiration to the canopy level for a given location. This assumption is also adopted in Donohue et al. (2013, 2017). We have made this assumption explicit in the manuscript (Line xxx).

It is also noted that Donohue et al. (2013, 3017) applied this equation at the same 5-year period as in the current study.

This reviewer did point out an important issue that this theory works better for undisturbed and mature vegetation but can be problematic for disturbed and immature vegetation (e.g., seedlings). However, the issue of vegetation age and disturbances is very complex and is well beyond the scope of this manuscript, especially considering that there are no global dataset monitoring vegetation age that we could use in our modelling. In the revised manuscript, we will point this issue out and discuss it.

R1C7: Notation: several symbols are defined differently from the publications they are taken from, creating some confusion. For example, mean rainfall depth is denoted by alpha (not beta) in Porporato et al. (2004); rooting depth is denoted by Z_r (not Z_e) in Guswa (2008); symbol beta is used in Guswa (2008) as well, but has a different meaning; many symbols are used to define evapotranspiration and potential evapotranspiration, and not all are clearly defined (E_{P_T}, E_T, E_{P_M}, E_P); stomatal conductance is generally denoted by g_s, not C_s; symbol theta is used for volumetric soil moisture (not water holding capacity). To summarize, for readers familiar with the literature, reading this manuscript can be difficult because of the different meaning of commonly used symbols.
Reply: Thanks for this comment; we will adjust the symbols used and make use that all symbols are clearly defined in the revised manuscript.

R1C8: L26: why "implicitly" - do you mean "explicitly"?
Reply: Yes, here should be explicitly; will revise in the revised manuscript.

R1C9: L31: "the resource availability gradient" suggests that this gradient has been presented before, but it is not.
Reply: We will change it to "a resource availability gradient" in the revised manuscript.

R1C10: L50: other recent works have discussed these issues, including Fatichi et al. (2016, www.pnas.org/cgi/doi/10.1073/pnas.1605036113).
Reply: We have read the suggested paper and agree that it is very relevant and will cite it appropriately in the revised manuscript.

R1C11: L61: please check spelling of BCP model author names.
Reply: Sorry for the typo. Will fix it in the revised manuscript.

R1C12: L67 and 69: are "model parameter" and "land surface parameter" indicating the same quantity?
Reply: Yes, they are the same parameter. We will use "model parameter" throughout in the revised manuscript to avoid potential misunderstandings.

R1C13: L81: this could be a good place for a summary of the research questions or aims of the work.
Reply: Will do in the revised manuscript.

R1C14: L96-97: just a comment - typically, ET is estimated from precipitation and runoff, since ET is the most difficult term in the catchment water balance to estimate; here the water balance is used to estimate Q, assuming the ET is known.
Reply: The Choudhary's formulation of the Budyko curve expresses actual ET as a function of P, PET and a model parameter. Then, the assumed steady-state water balance was used to calculate Q as a residual.

R1C5: L137: some words missing - e.g., "parameters"?
Reply: Oops! Will fix it in the revised manuscript.

R1C16: L139: but evaporation from the soil surface is neglected here (L91), so I am not sure I understand this statement.
Reply: We believe this reviewer misunderstood our approach. L91 reads "by taking soil surface resistance equal to zero". This mean that soil evaporation would occur at its potential rate.

R1C17: L147: I would define here symbols E_{P_M} and O.
Reply: We will define the symbol as suggested in the revised manuscript.

R1C18: L150: not clear how E_{P_M} differs from E_P.
Reply: E_{P_M} is the potential evapotranspiration that is only affected by meteorological conditions, while E_P is the potential evapotranspiration that is affected by both meteorological factors and CO2 concentration. With the increase of CO2, stomatal conductance decreases and LAI increases, both of which will affect potential evapotranspiration (actual evapotranspiration rate when water is not limiting). We will make the definition of E_P and E_{P_M} clear in the revised manuscript.

R1C19: L156: this sentence is hard to follow.
Reply: We will rephrase this sentence in the revised manuscript.

R1C20: L160: singular "affects".
Reply: Will do as suggested; thanks.

R1C21: Section 2.3: I would emphasize that this dataset covers experiments with artificially elevated CO2.
Reply: Will do as suggested; thanks.

R1C22: L210: how was beta calculated?
Reply: We will make this clear in the revised manuscript.

R1C23: L236: “differentially better” - meaning not clear.
Reply: We will rephrase it in the revised manuscript.

R1C24: L238: these statements are qualitative and no performance measure is provided to compare the two model variants.
Reply: We will add statistics to support this statement in the revised manuscript.

R1C25: L249: “. . .caused an increase of L” - in the remote sensing data or based on model predictions?
Reply: This is model prediction. The increased L reported here is only driven by the fertilization effect. We will explicitly state that this is a model prediction in the revised manuscript.

R1C26: L252: “L increase is found. . .” - in the remote sensing data or based on model predictions?
Reply: This is model prediction too. Please see our reply to R1C25.

R1C27: L265: suggested rewording “. . . shows a slight decrease in. . .”
Reply: Will do as suggested; thanks.

R1C28: L348: I am not sure how results here can guide climate model development.
Reply: We will reword this part in the revised manuscript.

R1C29: Figure 3: please check units of RMSE and mean bias in panel (b).
Reply: Oops! We will correct it in the revised manuscript.

R1C29: Figure 4: are the shown changes in L modelled or measured from remote sensing? Note that “but for each” in the caption is repeated.
Reply: This is modelled LAI change driven by the fertilization effect. Please see our reply to R1C25.

R1C30: Figure 6: I suspect L587-590 are not meant to be in the caption (they seem not relevant). I would also show error bars consistent with other plots – here they represent 1/10 of standard deviation, indicating that in fact the variance is extremely large.
Reply: We will correct the issues in the caption; please accept our apologies. The reviewer is correct that the variance among catchments is very large. This is because runoff and its changes differs quite a lot among catchments (Figures 3). However, it can be seen from Figure 6 that the variances are larger for climate change-induced runoff changes (precipitation and potential evapotranspiration). This is also reasonable, because the variances of P and PET changes can also be large among catchments. In direct contrast, the eCO2-induced runoff changes show a much smaller variance. We showed “1/10 of standard deviation” to that to better illustrate differences in means.

Thanks for your constructively critical review / assessment of our manuscript which provided the catalyst for improvement.

Anonymous Referee #2

R2C1: The manuscript by Yang et al. aims at quantifying the impact of physiological and structural vegetation adaptations induced by elevated atmospheric CO2 concentration (eCO2) on mean annual runoff (Q). The vegetation-mediated eCO2 effect on Q is complex and involved several processes with sometimes opposite effects. Also, the link of below-ground processes to eCO2 is still not entirely clear. For these reasons, the effect of eCO2 on Q is a source of uncertainty in simulation models. This paper uses an attribution framework, based on the previously applied BCP model, to quantify the net vegetation-mediated eCO2 effect on Q. This is a highly topical subject, the choice of methods seems appropriate and the inclusion of a link to below-ground processes constitutes a substantial novelty, which makes this manuscript of interest to HESS. However, my concerns relate to the presentation of the material: I find the manuscript difficult to follow and think that its value could be greatly increased by improving the description of methods. I therefore recommend a minor revision before the paper gets published.
Reply: Thanks for your favorable evaluation of our study. Your individual comments are replied to below.

R2C2: I find the presentation of the methods somewhat unclear and found it difficult to understand how the different methodical steps are linked together, particularly Sections 2.3 and 2.4. Are the responses of stomatal closure and L to eCO2 integrated in the BCP model? If so, please make the links explicit. If not, please clarify how these different steps work together in the attribution framework. Also, it seems to me that the step of extending the analysis from the study catchments (l. 196 states that the analysis is limited to those) to a global raster map (e.g. Fig. 7) is not described in sufficient detail in the Methods.
Reply: Thanks for your suggestion. We will document the methods more clearly in the revised manuscript.

R2C3: In the presentation of the results, it is not immediately clear if the Q-eCO2 response refers to the net effect of increased CO2 concentration on Q (through all the known effects on e.g. meteorological forcing, plant physiological and structural adaptations to CO2 and climate etc – this seems to be the case in the first paragraph of Section 3.3 and Fig. 5), or the net effect of eCO2-induced plant physiological and structural adaptations (this seems to be the case in the second paragraph of Section 3.3 and Fig. 6). Then again, in the first paragraph (l. 270 ff.) the authors discuss the relative importance of physiological and structural effects of eCO2 on vegetation before the corresponding evidence has been presented.

Reply: The reported effect does not include the CO2 effect on meteorological forcing but only for the CO2 effect on Q via vegetation feedbacks (plant physiological and structure responses). Thanks for your comment and we will make this clear in the revised manuscript.

R2C4: The authors conclude by stating that the analyses provide insightful guidance for the development of climate models. It would be helpful to describe how exactly the findings from this analysis can be used in climate model development. In general, if this is where the value of the paper lies, it would greatly benefit from connecting the different steps (methods and discussion) to the current state of research in climate and earth system modeling (including the significance of these feedbacks and their uncertainty for earth system modeling, e.g. Hickler et al. 2015 https://doi.org/10.1007/s40725-015-0014-8, Li et al. 2019 https://doi.org/10.5194/bg-15-6909-2018) . For example, how does the CO2 fertilization effect calculated in this paper compare to results obtained in modeling studies? How is the link between Ca and below-ground vegetation dynamics currently represented in models, and how might they benefit from the advances in this study?
Reply: Thanks for your suggestion. We will reword this part in the revised manuscript

R2C5: l. 111: Please indicate the values for root respiration and the Q10 parameters.
Reply: Will do as suggested; thanks.

R2C6: l. 142 ff: This is not necessarily the case. In the Guswa model, the relation of optimal rooting depth to P/EP is nonlinear and non-monotonic, with the greatest optimal depth calculated in conditions where water supply and demand are approximately equal.
Reply: The statement here discusses the potential CO_2 fertilization effect on rooting depth and not how rooting depth would respond to climate variations. To avoid confusion, we will make this explicit in our revision.

R2C7: Eq. 15: please define beta.
Reply: Beta was firstly introduced in line 101 and used in Eq. 2.

R2C8: l.161: what exactly does “residual” mean in this context?
Reply: Residual here refers to total dQ minus P-induced dQ plus E_{P_M}-induced dQ plus CO_2-induced dQ. In our parameterization, this other effect may include dQ caused by changes in rainfall intensity and climate-driven vegetation changes. We will make this clear in the revised manuscript.

R2C9: Eq. 16: What are the units of S_Q_to_eCO2?
Reply: Apologies for our oversight; the units of $S_{Q\_to\_CO2}$ in Eq. 16 is mm yr^{-1} ppm^{-1}. We will add the units in the revised manuscript.

R2C10: l. 250: This average value by itself is not very informative, I suggest

characterizing the distribution (mode(s) and range) in more detail (including a discussion of Fig. 4 b).
Reply: Will do as suggested; thanks.

R2C11: l. 257 "has resulted": I suggest making it clearer that this statement describes simulation results, rather than observations (as I understand it).
Reply: Good point and yes we will do as suggested as trends from simulation have less 'scientific weight' than observed trends; thanks. We will say something like "the simulated resulted showed" or something similar.

R2C12: l.288: did you mean "other factors including"?
Reply: Oops! Will revise as suggested.

R2C13: l. 320: which mechanism?
Reply: We will reword this in the revised manuscript to avoid it being potentially ambiguous.

R2C14: l.337: I am not sure if the word "exaggerate" corresponds to the idea expressed by the authors. Maybe "exacerbate"?
Reply: Sorry for the typo. Will be revise as suggested. Thanks for your careful review; its appreciated.

R2C15: l. 340 "This suggests that the structural response. . ." This causal link is not immediately clear to me, please clarify
Reply: We will reword this part to make it clear in the revised manuscript. Basically, the structural response of vegetation to eCO_2 decreases with the increase of leaf area index, so with the increases of leaf area index in future climate scenarios (due to higher P and CO_2), the response of vegetation structure (e.g., leaf area index) to elevated CO_2 will eventually decreases.

R2C16: Fig. 4 a,b,d,e: To avoid any confusion I think it is important to make clear that the data shown are the results of simulations, and not (as I understand it) based on observations.
Reply: Thanks for the suggestion. Will do as suggested, and please see our reply to R2C11 above.

R2C17: Fig. 5: What exactly is meant by "Q change induced by eCO2" (see my comment in the 3rd paragraph)?
Reply: We will make it clear that this eCO_2 effect only refers to the eCO_2-induced effects via vegetation feedbacks on Q in the revised manuscript.

R2C18: Fig. 6: The size of the error bars representing 1/10 suggests a great variability of these quantities among the different catchments. Consider using an alternative visualization method (e.g. boxplots or kernel density plots).

Reply: Good point; we will find a better way to present this result in the revised manuscript. Please see our reply to R1C30.

R2C19: Fig 6: some sentence of the caption refer to elements that I cannot see (viewing the PDF in Chrome on Windows): values in parenthesis; vertical grey dashed line.
Reply: We will correct this issue in the revised manuscript.

Thanks very much for your careful and detailed review which will result in our manuscript improving.

Anonymous Referee #3

R3C1: The authors explore past runoff trends over undisturbed catchments and globally. Using an analytical framework, they attribute runoff trends to climate and vegetation influences along a resource availability index. The impact of CO2-induced vegetation changes on runoff has remained highly uncertain and as such, this study is a valuable contribution to the literature and well suited to HESS.
Reply: Thanks for your encouraging and constructive comments. We reply to your individual comments below.

R3C2: Whilst I find this study interesting, it is a shame it does not go further in quantifying the CO2-induced vegetation changes on Q. In particular, the authors mention the inclusion of CO2-induced rooting depth changes as a key novel aspect. However, whilst the authors quantify in detail the influence of CO2 on the individual above- and below-ground vegetation processes, it is not shown how these in turn affect Q. For Q, only the bulk CO2 response is presented if my reading of the results is correct. A number of studies already exist on the bulk CO2 responses and/or separating the effects of stomatal closure and LAI on Q (although I appreciate a new modelling framework is introduced here). Here it would have been interesting to know how Ze specifically changes Q. I think the results suggest the influence of rooting depth changes are minimal but this is glossed over in the discussion.
Reply: This is a very good suggestion; thanks. However, in our framework, the impacts of eCO_2 on vegetation structures (both LAI and rooting depth) are effectively calculated simultaneously, making it difficult to examine the LAI-induced and rooting depth-induced runoff changes separately. Specifically, eCO_2-induced rooting depth change is parameterized through changes in WUE and potential transpiration, and the eCO_2-induced potential transpiration change is caused by eCO_2-induced changes in stomatal conductance and LAI. Hence, the framework developed here does not allow a separate evaluation of LAI and rooting depth effects. Following your suggestion, we will perform new analysis to examine how the physiological and structural responses of vegetation to eCO_2 affect runoff separately.

R3C3: I would also hope more clarity on how parameter n is determined. The current explanation is not sufficient, including what data were used. The methods should also be revised for clarity, reading the results it becomes unclear what was quantified using the analytical framework vs other methods (e.g. stomatal closure and L responses). Perhaps a summary of the steps at the start of Methods would help the reader.
Reply: Thanks for the suggestion. We will further clarify the methods in the revised manuscript. Please see our response to R2C2, above.

R3C4: Title: The study period doesn't cover the last three decades.
Reply: to be explicit regarding the study period in the title will change it to be "Low and contrasting impacts of vegetation CO_2 fertilization on global terrestrial runoff over 1982-2010: Accounting for above- and below-ground vegetation-CO_2 effects"

R3C5: L23-26: This sentence could be written more clearly.
Reply: We will rephrase this sentence to make it more clear.

R3C6: L30-34: This sentence should also be broken up into two for clarity.
Reply: Will do as suggested; thanks.

R3C7: L34: highlights -> highlight
Reply: Will do as suggested; thanks.

R3C8: L38: Suggest replacing "becoming" with "and representing"
Reply: Will do as suggested; thanks.

R3C9: L44: I would suggest Donohue is not an appropriate reference here, it is not a leaf-scale study.
Reply: We will delete this reference and insert an appropriate leaf-scale study citation here.

R3C10: L50: I think the authors need to unpack this sentence a little. Many of these studies look at the net response on Q so I'm not sure what the authors mean by "different aspects"? I would argue the main reason for the discrepancies across studies is due to the different processes and assumptions included in the models. Also Ukkola et al. is not a modelling study but based on observations (similarly Trancoso et al. 2017 which should also be cited for an observational analysis). Some model evaluation has also been conducted specifically for CO2 impacts (e.g. Ukkola et al. 2016, Environmental Research Letters for a DGVM and multiple FACE papers), here the evaluation seems to be limited to the overall Q trends which is not new. A more accurate statement would be that observational and evaluation studies for CO2 effects remain limited, particularly at the regional to global scale.
Reply: Thanks for the suggestion. We will rephrase this part in the revised manuscript as suggested.

R3C11: L53, L265: please fix grammar.
Reply: Will do; thanks.

R3C12: L61: should be Budyko-Choudhury-Porporato
Reply: Oops! Will revised as suggested; thanks.

R3C13: L94: I think Milly and Dunne actually found that the energy-only PET best produced non-water-stressed ET from climate models (their Figure 3, associated text and conclusions). Also climate models do not simulate potential evapotranspiration so perhaps best to avoid that terminology here?
Reply: Milly and Dunne (2016) used the two-source Shuttleworth-Wallace model to examine the effect of eCO_2 on vegetation stomatal conductance and the consequent impact on potential evapotranspiration (the actual evapotranspiration when water is not limiting) (their Figure 2 and Figure 3). The calculation of PET using the two-source Shuttleworth-Wallace model in our study follows Milly and Dunne (2016; https://doi.org/10.1038/nclimate3046). Yang et al. (2019; https://doi.org/10.1038/s41558-018-0361-0) further demonstrated that why Milly and Dunne found energy-only PET works well is because that the warming-induced PET increases are almost entirely offset by the eCO_2-induced PET decreases. By using the two-source Shuttleworth-Wallace model, we are able to explicitly consider the impacts of LAI and conductance changes on PET.

We agree that the climate models do not simulate PET. We will rephrase this statement in the revised manuscript.

R3C14: L103: Could you be more specific here? Taking what data?
Reply: The data obtained from the authors of that paper is their numerical solutions of the corresponding E/P for every 0.1 increment in PET/P for six separate γ curves. By numerically solving the Choudhary's formulation of the Budyko curve, we determined the values of the Budyko parameter (n) that correspond to the E/P values of each of the six γ curves. We then pooled all $n - \gamma$ pairs together to derived a simple relationship between them (Eq. 2 in the manuscript).

We will make this clear in the revised manuscript.

R3C15: L111: Not clear to me how potential transpiration is determined?
Reply: Potential transpiration was determined by applying the two-source Shuttleworth-Wallace model while assuming a non-water-limiting condition. This two-source approach allows evaporation from soil and transpiration from vegetation to be calculated separately. We will extend the description of potential evapotranspiration calculations to make it clear in the revised manuscript.

R3C16: L119: I don't see where the Earth System Models are described? Also why were ESMs used rather than something more observationally constrained? Given such a short time period is taken and coupled models have their own interannual variability, taking a mean across models over such a short time period is likely to be spurious. Why wasn't observationally-driven products used, e.g. GLEAM or the TRENDY ensemble? These are of course also models but at least driven by observed meteorology

Reply: We are sorry that we missed the description of the ecosystem models used in this study. To obtain a spatial pattern of *WUE*, global monthly GPP and E_T estimates over 1982-1985 were obtained from 8 ecosystem models from MsTMIP (Multi-scale Synthesis and Terrestrial Model Intercomparison Project; *Huntzinger et al*. [2013]), including: (i) CLM [*Mao et al.*, 2012]; (ii) CLM4-VIC [*Li et al.*, 2011]; (iii) ISAM [*Jain et al.*, 1996]; (iv) TRIPLEX [*Peng et al.*, 2002]; (v) LPJ-wsl [*Sitch et al.*, 2003]; (vi) ORCHIDEE-LSCE [*Krinner et al.*, 2005]; (vii) SiBCASA [*Schaefer et al.*, 2008]; and (viii) VISIT [*Ito*, 2010]. MsTMIP is a model comparison program, which is similar to TRENDY. All models participated in MsTMIP are forced by observed meteorological variables. In fact, the same modelling outputs of these ecosystem models contributed to MsTMIP and TRENDY at the same time. We will add relevant descriptions in the revised manuscript.

R3C17: L138-139: Why do these quantities impact Ep? Most PET estimators are mainly atmosphere-driven so if this is not the case with Shuttleworth and Wallace, more details on its calculation need to be provided for clarity.

Reply: The Shuttleworth-Wallace model is a two-source evapotranspiration model, which calculates soil evaporation and plant transpiration separately. To be able to distinguish the two flux sources, vegetation conditions are needed in the model. We will extend the descriptions of the Shuttleworth-Wallace model and how potential evapotranspiration and its components (potential evaporation and potential transpiration) are calculated in the revised manuscript.

R3C18: Equation 12: should the notation be f() instead of g()?

Reply: Using f() and g() indicate that the two functional relationships are different.

R3C19: L218: Which years were used?

Reply: The year of 2001 was used. We will add this information in the revised manuscript.

R3C20: L220: You should provide the name of the dataset (i.e. ISLSCP etc.)

Reply: Will do as suggested; thanks.

R3C21: Figure 1: White regions in the map that do not match the colour scale (e.g. Greenland). Should say if/why these were masked out

Reply: Will do as suggested; thanks.

R3C22: L232, L234: missing full stop

Reply: Oops! Will correct it in the revised manuscript.

R3C23: L243: Please avoid using brackets like this, it is very hard to read. Suggest: with the largest Cs reduction found in C4 crops and lowest in shrubs.
Reply: Will revise as suggested; thanks.

R3C24: L249: How was the Ca effect on L estimated? I'm assuming using equation 18 but it has two factors influencing L (Ca and v)
Reply: This reviewer was correct that both Ca and *v* are considered in our estimation of how *L* respond to changes in Ca. The reason that L increases with elevated CO_2 is because that WUE increases with elevated CO_2. However, the response of WUE to elevated CO_2 is also mediated by changes in *v*. In our modelling framework, we firstly estimate how WUE changes and then how *L* respond to changes in WUE. In the revised manuscript, we will try to rephrase relevant text to avoid potential misunderstandings.

R3C25: Figure 5: Would be useful to see the spatial distribution of catchment trends. Suggest adding a map of the catchments eCO2-induced trends as an additional panel
Reply: Will do as suggested; thanks.

R3C26: Figure 6: Last panel please adjust scale to show full error bars. Also please check caption, from L587 it mentions numbers that I don't see presented in the figure
Reply: Will do as suggested; thanks.

R3C27: L279: I'm confused why this result differs from the number on L269 and how the changes in Q described here differ from the previous paragraph?
Reply: Oops! Sorry for the typo. The change should be -2.3 mm yr^{-1}. We will correct this number on L269 in the revised manuscript.

R3C28: L286: I'm also confused that you have suddenly moved to global results (Fig 7). In the methods, you state that the analysis is restricted to the ~2000 catchments (L195). The text doesn't also make this transition obvious.
Reply: Sorry about that, we will add necessary information in the Methods section and revise relevant text in the Results section to make this transition more smoothly.

R3C29: L292: Given alpha is determined from LAI, surely low-alpha regions can be either dry or cold?
Reply: Yes, this reviewer was correct.

R3C30: L348: How exactly can this framework guide model development? Firstly, the results from this study are very much in line with existing studies so no particularly novel insights are revealed. And secondly, how is this framework to help climate model development exactly? And finally, this is ultimately simply another model result. Overall this feels like a bit of a throw-away statement to try and boost the value of paper.

Reply: We will reword this part to avoid overstatement.

R3C31: L351: Are all the datasets publicly available?
Reply: Yes, they are publicly available. We will make sure that all data sources are explicitly documented in the revised manuscript.

Finally, thanks very much for your review which has allowed us to improve the science in our manuscript.