Response to reviewers

Response in blue

Reviewer 1:

The manuscript by Stacey et al. examines how vegetation responses to rising atmospheric CO_2 and climate change affect global water scarcity projections using the JULES model. The authors assess how physiological (stomatal closure) and structural (leaf area and vegetation cover) changes influence future water supply. The results suggest CO_2-induced stomatal closure increases water availability, potentially alleviating scarcity, particularly in tropical regions, while vegetation expansion may worsen scarcity in some semi-arid and arid areas. Overall, this study makes a valuable contribution to the literature on vegetation-climate-hydrology interactions, and I appreciate the authors' analysis and findings.

Thank you for the summary and kind comments. We greatly appreciate your time and effort in reviewing the paper.

Broader suggestions and concerns

1. Since this study relies on a single model, a more detailed description of key model processes would help readers better understand potential limitations. While a full breakdown of all JULES equations is unnecessary, some critical details are missing. In particular, a brief explanation of the coupled canopy conductance and photosynthesis model from Cox et al. (1998) would be helpful, as the results strongly depend on this component. (For instance, does this simplification overestimate the physiological impacts?) The discussion on limitations (L553-564) mentions issues with stomatal conductance parameterization; rather than only stating this as a limitation, it would be more helpful to discuss how it affects the key findings.

A description of the relevant stomatal conductance and photosynthesis equations will be added as Appendix B:

"JULES uses a coupled canopy conductance and photosynthesis model (Cox et al., 1998) where stomatal conductance to water vapour g_s (m s^{-1}) is based on:

$$g_s = -1.6A\frac{RT^*}{c_i - c_a} \tag{1}$$

where *A is* the net photosynthetic rate (mol CO_2 m^{-2} s^{-1}), *R* is the universal gas constant (J K^{-1} mol^{-1}), *T** is the leaf surface temperature (K), c_i the internal CO_2 partial pressure (Pa), c_c the leaf surface CO_2 partial pressure (Pa) and factor of 1.6 accounting for molecular diffusivity differences between water and CO_2. Vapour deficit at the leaf surface (D, kg kg^{-1}) affects stomatal conductance through the gradient between c_a and c_i is based on the equation by Jacobs (1994):

$$\frac{c_i - \Gamma^*}{c_a - \Gamma^*} = f_0\left(1 - \frac{D}{D_{crit}}\right) \tag{2}$$

where Γ^* is the photorespiration compensation point (Pa) and D_{crit} and f_0 are PFT-specific calibration parameters, which are directly related to the parameters from the Leuning (1995) model (for details see Cox et al., 1998). Potential non-stressed leaf level photosynthesis is calculated in JULES using the C3 and C4 photosynthesis models of Collatz et al. (1991) and Collatz et al. (1992) respectively."

You raise a good point regarding how this may affect our findings. This is a difficult one to answer on the global scale, due to lack of observational data. We will add the following text to the discussion to explain this.

"Another source of uncertainty lies in the accuracy of the parameterisation schemes used in JULES to represent hydrological and biophysical processes. For example, the stomatal conductance scheme simplifies a complex process which varies across species, ecosystems, and climates (Norby and Zak, 2011). Evaluating the accuracy of such parameterisation schemes on a global scale is a major challenge, primarily due to limited observational data. Experiments such as Free Air CO_2 Enrichment (FACE) have provided valuable insights into plant physiological responses to elevated CO_2, which are crucial for land surface model developments, but are currently at a small number of point locations. Model-data comparison studies suggest mixed performance at simulating CO_2 effects on water-use efficiency, with models performing well for some sites and species but poorly for others (De Kauwe et al., 2013; Walker et al., 2014). Although JULES has not yet been comprehensively assessed in these studies, its behaviour expected to be broadly comparable to the models evaluated. Furthermore, some studies suggest that the magnitude of the CO_2-effect on river runoff records in JULES is reasonable within uncertainty bounds (Gedney et al., 2006, 2014)."

2. In some cases, such as in Section 3.1, the results come across as relatively intuitive. While I understand that this section serves to introduce the different simulations and their effects on vegetation and water cycle variables, I would argue that the main novelty of the study lies in the quantification and analysis of the Water Scarcity Index (WSI). For that reason, Section 3.1 could be streamlined or more tightly focused, allowing the reader's attention to shift more quickly toward the WSI results and their broader implications.

Thank you for the suggestion - section 3.1 will be condensed.

3. The discussion section contains many well-acknowledged limitations, which is appreciated. However, rather than listing them in loosely connected paragraphs, they should be structured to explicitly discuss how each limitation affects the main results. Some parts could also be shortened and streamlined, for instance, the justification for using runoff as a proxy for water availability is reasonable in the context of its role as a water supply source. In my opinion this could be shortened to one sentence.

Again, thank you for the advice. The discussion section will be condensed, and the limitations sections restructured to emphasise the effect on the results.

4. While this study is an idealized experiment rather than a model performance evaluation, I think it's important to acknowledge that non-linear biases can emerge when driving JULES with biased GCM inputs. Even though the input fields are identical across simulations, their biases may interact differently with the model's internal processes. This contradicts your statement in L569-571: „Given that our study focuses on the influence of plant responses – comparing the differences from two simulations with similar biases - HADGEM2-ES and this version of JULES are deemed suitable for our purposes, despite the runoff biases." A brief discussion on how these biases might influence the results would improve transparency.

Thank you for this query and suggestions. The driving climate model data is already bias-corrected, and we expect that the relatively runoff biases we mention are due to JULES itself rather than the driving GCM data. We also consider these biases to be relatively small. Nevertheless, we agree that their potential implications should be discussed, although we are not sure it is possible to actually quantify these. We will add the following to the discussion section:

"The driving climate data from HadGEM2-ES used to force JULES in this study has been bias corrected following the ISIMIP2b protocol, although some residual biases remain in the version of JULES used in this study (Mathison et al., 2023). For the period 1980–2006, their evaluation indicates negligible runoff biases for most river basins. However, slight underestimations in runoff are found in China and the northern high latitudes, leading to overestimations in WSI for the present day. This could mean that the potential for CO_2 effects to bring WSI into a less severe state in this region may be underestimated. Meanwhile, slight overestimations in runoff in eastern USA may result in underestimations in WSI, which could mean that the potential for climate change to lead to higher WSI even in the presence of CO_2 effects may be underestimated. To assess the influence of plant responses, we take the difference between two simulations that share similar biases, which helps to minimise overall biases. Non-linear biases may persist when comparing simulations with differing plant responses, but these are likely to be relatively small compared with other uncertainties inherent in such modelling studies."

Minor and Editorial Comments

L16: Simplifying and breaking this sentence into clearer segments would improve its readability and ensure that the statistical comparisons are more immediately accessible to the reader.

Thanks for the suggestion. We have hopefully made this sentence clearer and more digestible – see below - but welcome any further recommendations.

"For the future period (2076–2095), incorporating all plant responses lowers global median WSI by approximately 12%. Furthermore, across 291 global river basins, median WSI is lower (by 10-70%) in 138 basins, home to 80% of the global population, and higher (by 10-60%) in 11 basins, affecting 0.2% of the population."

Section 2.3 (Data): This is more of a stylistic suggestion, but I found this section somewhat fragmented and lacking in flow. A more cohesive, narrative-style format could improve readability. Consider merging some of the sub-subsections to reduce fragmentation, and integrating the definitions of key terminology directly into the text rather than listing them separately in bullet points. This would help create a smoother and more engaging structure.

We will rearrange the Methods section. Thank you for the suggestion.

Figures: The manuscript contains many figures, and some could be streamlined. In my opinion you could merge Figure 1 and Figure 3 and move it to the supplementary materials or place it in the methods section, clarifying that it shows HadGEM2-ES input data to JULES. In Figure 2, adding simulation abbreviations (e.g., S2 - S1 CLIM: VEG) in the second legend for isolated factors could improve clarity.

Figure 1 and 3 will be moved to the Methods section as Figure 1A and 1B. The additional abbreviations will be added to Figure 2 and also Figure 4.

Tables 1a and 1b: These were difficult to interpret at first glance. It might improve clarity to describe Table 1a in more detail in the text and move it to the supplementary materials. For me Table 1b was the more informative one in explaining the mechanisms, so I suggest keeping the focus there.

Table 1A will be added as an Appendix and an explanation on each simulation added to the Methods section:

"Simulation 1 (S1) CLIM: STOM has fixed vegetation distribution and physiological forcing and is closest to a typical hydrology study. Simulation 2 (S2) CLIM: STOM+VEG has fixed physiological forcing only, with dynamic vegetation distribution included. Simulation 3 (S3) CLIM+CO2: STOM has fixed dynamic vegetation distribution only, with physiological forcing switched on. Simulation 4 (S4) CLIM+CO2: STOM+VEG includes all physiological and structural responses to CO_2 and climate and is closest to fully coupled Earth System Model. Different "isolated factors" are combined by taking the differences between simulations. CLIM: VEG (S2 – S1) represents climate effects on vegetation coverage/LAI. CO2: STOM (S3 – S1) represents CO_2 effects on stomata. CO2: STOM+VEG (S4 – S2) represents CO_2 effects on stomata and vegetation coverage/LAI. CO2: STOM & CLIM+CO2: VEG (S4 – S1) represents CO2 effects on stomata and climate plus CO2 effects on vegetation coverage and LAI, and indicates the differences between a typical offline hydrology study and the fully coupled ESM."

L263: Subject-verb agreement: "Climate-induced changes in vegetation distribution **drive** LAI decrease," not "drives."

Thank you - this will be changed.

L293: The statement: "The climate-induced changes in runoff from the present (2006-2025) to the future (2076-2095) largely align with shifts in precipitation (Fig. 3; right), varying greatly in both magnitude and direction (Fig. 4a,c)." seems contradictory. If they largely align, how can they vary greatly in direction?

Thanks for spotting this. Hopefully this is clearer now:

"Climate-driven runoff changes from the present (2006–2025) to the future (2076–2095) period are highly variable in both magnitude and direction (Fig. 3a,c) generally aligning with precipitation changes (Fig. 1b; right)."

L308 – 310: You state that the runoff decreases are driven by CO2 induced LAI increases, but this is hard to see on the map and that raises the question of how large and significant the impact really is.

Thanks for the suggestion. We have now hopefully made this clearer – see text below. We agree that the decreases are relatively much smaller than increases, but the plot does show light beige colours across vast parts of the globe, representing very small decreases, which as now explained in the text, could be impactful in some areas. The following will be added to the results section to clarify this.

"However, CO_2-induced LAI increases appear to drive slight runoff reductions across some areas, particularly in semi-arid and arid climates such as the Middle East and western USA (Fig. 3i,l). Although considerably smaller than the increases, these **small runoff reductions could cause significant impacts in already water-stressed areas."**

General formatting: Check for missing spaces between words, unclosed brackets, and misplaced semicolons throughout the text (e.g., L293).

We will check the text. Thank you for pointing these out.