List of changes made

Reviewer 1

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?)

A description of the relevant stomatal conductance and photosynthesis equations has been added to Appendix B:

"Stomatal Conductance Scheme used in JULES

The version of JULES used in this study 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:

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

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):

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

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."

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.

The following text has been added to the discussion:

"JULES uses parameterisation schemes to represent hydrological and biophysical processes. For example, the stomatal conductance scheme (see Appendix B) 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 CO2 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 is expected to be broadly comparable to the models evaluated. Furthermore, some studies suggest that the magnitude of the CO2-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.

Section 3.1 has been 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.

The discussion section has been 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.

We have added the following to the discussion section:

"The HadGEM2-ES climate data driving JULES in this study has been bias corrected following the ISIMIP2b protocol, although some residual biases remain in the JULES version used (Mathison et al., 2023). For 1980–2006, their evaluation indicates negligible runoff biases for most river basins. However, slight runoff underestimations in China and the northern high latitudes translate to overestimations in WSI in our study. Conversely, slight runoff overestimations in eastern USA translate to underestimations in WSI in our study. To assess the influence of plant responses, we take the difference between two simulations that share similar biases, and thus this minimises overall biases. Non-linear biases may persist when comparing simulations with differing plant responses, but these are expected 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.

The following text has been added to the abstract:

"For the period 2076–2095, incorporating all plant responses to CO_2 and climate change reduces global median WSI by approximately 12%. Furthermore, across 291 river basins, 138 basins show lower median WSI (by 10-70%), representing 80% of the global population, while 11 basins show higher WSI (by 10-60%), representing 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.

The Methods section has been rearranged.

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 have been moved to the Methods section as Figure 1A and 1B. The additional abbreviations has been added to Figure 2 and 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 has been as Appendix 1. An explanation on each simulation has been added to the Methods section:

"Simulation 1 (S1) CLIM: STOM includes climate-induced stomatal changes only and is closest to a typical hydrology study with fixed vegetation coverage/LAI and physiological forcing. Simulation 2 (S2) CLIM: STOM+VEG includes climate-induced stomatal and vegetation coverage/LAI responses. Simulation 3 (S3) CLIM+CO2: STOM includes CO2- and climate-induced stomatal responses. Simulation 4 (S4) CLIM+CO2: STOM+VEG includes all CO2- and climate-induced stomatal and vegetation coverage/LAI responses 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 CO2 effects on stomata. CO2: STOM+VEG (S4 – S2) represents CO2 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/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."

This sentence has been modified.

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?

Now changed to:

"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) 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.

The following has been added to the results section:

"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).

All text has been checked thoroughly and errors corrected.

Reviewer 2

General comments

The authors focus on a single land-surface model with prescribed atmosphere. The authors could briefly discuss how the findings of these manuscript fit into the broader context. How can the findings from this single model study be applied to other models?

The following has been added to the Discussion section:

"Since only one land surface model was assessed in this study, further work could test additional models to examine the sensitivity of the results. We would anticipate broadly consistent outcomes across LSMs when testing the stomatal response, as many LSMs employ stomatal conductance parameterisation schemes derived from similar formulations. However, we would expect a wider spread of responses in LSMs when testing vegetation structural responses, given that different Dynamic Global Vegetation Models apply alternative approaches to dynamic vegetation compared with the TRIFFID scheme used in JULES (Sitch et al., 2008)."

The following has been added to the Conclusion section:

"Our findings highlight the need to incorporate vegetation dynamics, primarily physiological forcing, into hydrology models to improve the robustness of water scarcity assessments."

The stomatal conductance model used in JULES is the Leuning one, and it would benefit readers to include a brief explanation on the variables represented in the Leuning model and how it differs from other models (e.g., Ball-Berry, Medlyn, etc.) in formulation and impact on water scarcity.

The following has been added as Appendix B:

"Stomatal Conductance Scheme used in JULES

The version of JULES used in this study uses a coupled canopy conductance and photosynthesis model (Cox et al., 1998) where stomatal conductance to water vapour gs (m s−1) is based on:

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

where A is the net photosynthetic rate (mol CO2 m−2 s−1), R is the universal gas constant (J K−1 mol−1), T* is the leaf surface temperature (K), ci the internal CO2 partial pressure (Pa), cc the leaf surface CO2 partial pressure (Pa) and factor of 1.6 accounting for molecular diffusivity differences between water and CO2. Vapour deficit at the leaf surface (D, kg kg-1) affects stomatal conductance through the gradient between ca and ci is based on the equation by Jacobs (1994):

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

where Γ* is the photorespiration compensation point (Pa) and Dcrit and f0 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.

The Jacobs formulation is a simplified version of the Leuning (1995) model, which in turn is based on the (Ball et al., 1987) model but depends on humidity deficit at the leaf surface instead of relative humidity."

The use of physiological and structural response in addition to stomatal and leaf area responses in the introduction is confusing to the reader. It could be helpful to directly link physiological and structural responses to stomatal and leaf area responses. I understand that the authors define "physiological forcing" in the terminology, but the manuscript would benefit if the distinction is made more clear earlier. Additionally, does "vegetation distribution change" refer the structural response?

All "vegetation *distribution* changes" have been changed to "vegetation *structural* changes" and the following has been added to the beginning of the introduction:

"Throughout this manuscript, we use the following terminology:

- **Stomatal response** – changes in stomatal conductance (i.e., stomatal opening and closure)
- **Structural response** – changes in vegetation structure (leaf area and canopy coverage)
- **Physiological response (or forcing)** – CO_2-induced changes encompassing both stomatal behaviour and vegetation structure (leaf area and canopy coverage)"

Minor specific comments

Table 1a and b were very helpful in understanding the different simulations. An additional line separating calculations vs simulations would help the reader interpret the table, though I do appreciate the bolding attempts for that.

A line has been added to separate the simulations and calculations in Table 1.

Figure 1 could benefit from removal of the legend since data from only one simulation was used to create the plots. Instead, an overall title (e.g., 'Annual mean timeseries from S1. CLIM: STOM). The removal of the legend will also allow the plots to be larger.

Legends have been removed from Figure 1.

Runoff as a proxy for water supply intuitively makes sense, but has not been explicitly related in the introduction or methodology section. A sentence for why runoff is used as a proxy in the methodology section would be helpful to the reader. Additionally, surface and sub-surface runoff are analysed, and a distinction and description for the two and what they can tell us would be helpful.

The following has been added to the Methods section:

"Runoff is used as a proxy for water supply because it represents the fraction of precipitation that is not lost to evapotranspiration and is therefore available for water resources. Total runoff is taken from the output of the simulations and includes both surface (overland flow generated when precipitation exceeds infiltration capacity or soil saturation) and sub-surface runoff (lateral drainage through the soil column) at the grid box scale."

L360: The "decreases" in the last two sentences are not specified and confusing to the reader.

Thank you for pointing this out. We have amended the text which hopefully makes it clearer:

"When CO_2-induced vegetation structural changes are also included in the CO2: STOM+VEG simulations (Fig. 6h,i), overall supply still increases, as the supply reductions associated with CO_2-driven structural vegetation expansion are relatively small and less apparent in the plots. However, these modest reductions translate into substantial increases in WSI in arid regions such as the Middle East and northern and southern Africa (Fig. 6i), where baseline water supply is already low."

The right plot in Fig. 8 "% difference" could state "% difference of median WSI" for the reader's ease of reading.

Figure 7 (previously Figure 8) x-axis now titled "% Difference in Median WSI"

Technical corrections

L151: extra space after period: "bias correction ."

L231: Extra space between "SSP2" and "from"

L264: extra ";" after "Fig. 2n"

L416: Extra ",)" after "(Fig. 9d)"

Generally, double check for extra spaces, missing spaces, commas, and semicolons throughout the manuscript.

The corrections have been made and manuscript carefully checked.