**RC1**

## 1 General comments

The paper provides a rich global-scale analysis of effects of land use and land cover change (LULCC) upon the partitioning of blue and green water flows, using a LUMIP model ensemble in different setups (focused on a 'sustainable LULCC' scenario). The analysis appears to be well done and plausibly interpreted, with a quite detailed examination of process interactions, regional features, and model uncertainties. That said, the Results and Discussion are rather dense with much information, parts of which could be portrayed in a somewhat more accessible way. Below I provide some comments on how to possibly achieve this. Overall, may comments are relatively minor and could be addressed mainly by some clarifications and restructurings.

We thank the reviewer for the thorough and constructive assessment. We are glad the analysis was found to be well executed and plausibly interpreted. We agree that the Results and Discussion are currently dense, and several of the revisions below are aimed specifically at improving accessibility. Our point-by-point responses are given below in blue.

## 2 Specific comments

The Abstract should clarify that you mainly analyze transpiration (as a representative of green water) rather than total evapotranspiration. While outputs for evaporation appear to have been analyzed also, I do miss a short discussion of whether a full inclusion of evapotranspiration (not just transpiration) in the ratio would produce relevantly different results.

We agree. We will state explicitly in the Abstract that the green-water flow is represented by transpiration rather than total evapotranspiration, and add a short Discussion paragraph on the consequences: transpiration isolates the productive, vegetation-controlled flux, whereas total ET also includes soil evaporation and interception that can vary independently of vegetation. The component fields $\Delta E$, $\Delta ET$ and $\Delta E_{\mathrm{t}}$ are already provided (Figs. S8, S10, S11), and we will add a supplementary figure indicating where the choice of green-water flux matters most.

Please state whether (and how) the direct effects of increasing carbon dioxide concentrations upon vegetation productivity / coverage and upon transpiration is considered in the models. No need to separately quantify the particular contributions of these effects here, but at least some words on this would be helpful (in the Discussion); i.e. is transpiration suppressed a lot under SSP3-7 / high CO2 concentrations. Relatedly, the UKESM model simulates "dynamic PFT competition" while in other models vegetation distribution is prescribed (according to Table 1). Does this have implications for results from that particular model?

We will expand the Discussion accordingly. Because the simulations are concentration-driven and the paired land-use runs within each background share the same $CO_2$ pathway (S1L1/S1L3 and S3L1/S3L3), the direct CO2 physiological effect is controlled for in the paired land-use differences and therefore does not directly drive $\Delta$LULCC; any remaining influence would arise through $CO_2$ $\times$ land-use interactions. We find this interaction to be weak for transpiration: the land-use transpiration response is nearly identical under the low-$CO_2$ (SSP1) and high-$CO_2$ (SSP3) backgrounds (spatial correlation $r \approx 0.90$, with similar overall magnitudes), so transpiration is not systematically suppressed at high $CO_2$ in our isolated signal. For UKESM1, we will note that its TRIFFID dynamic vegetation (prognostic PFT fractions within natural/managed units) can reinforce transpiration responses, contributing to its tendency toward larger green-water shares.

Sections 3.1 and 3.2 contain a lot of information with many acronyms etc. Is it possible to provide a table listing the model experiments (that could also be used to remind the reader of the different setups analyzed) and key results for each, such as the correlations, global areas affected by change XY, or something similar? Then the table could be pointed to for these results rather than mentioning them all in the text.

We will add a summary table that (i) lists the LUMIP experiments and the four derived simulations (S1L1, S3L3, S1L3, S3L1) with a one-line reminder of each setup, and (ii) summarises the key quantitative results (e.g. $I_{\mathrm{BGWS}}$, sign-agreement fractions, characteristic correlations). The text in Sects. 3.1–3.2 will then point to this table rather than repeating the numbers, reducing acronym density.

Besides, if there is need for shortening the length of the text, section 3.3 is a candidate; it is a regional zoom of the analysis providing much detail, maybe not all of it required to get the main messages across.

We will shorten Sect. 3.3 by condensing the model-by-model description and moving supporting detail to the Supplement, while retaining the key message that the same forcing yields divergent partitioning responses across models. The schematic figure of uncertainty sources (requested by Reviewer 2) will help replace several descriptive passages with a more compact synthesis.

At the beginning of the Discussion, the key results could be highlighted again as bullet points, and then – which I would definitely recommend – the Discussion should have some subtitles for each of these points (like green water shares, teleconnections, uncertainties, limitations). This would be another, easy-to-implement way to highlight once more the main results and to guide the reader through the large amount of information.

We will restructure the Discussion to open with a short summary of the main findings and add subsections, e.g. on vegetation response, precipitation feedbacks, uncertainty, and limitations.