Referee comments are shown in black. Our responses in bold and yellow.

RC2: 'Comment on hess-2023-255', Anonymous Referee #2, 20-12-2023

General comment:

Sánchez-Dávila and co-authors present an exhaustive modelling exercise to quantify green and blue water fluxes across Spanish forests (excluding the Canary Islands). These estimations are based on a very detailed forest ecosystem model, which has been evaluated against ecophysiological data at the plot level and against forest dynamics in Spain, with good results. In this paper, the results on blue water (BW) and %BW and the patterns associated with climate and functional groups are somewhat expected, but nevertheless valuable, as they had not been examined with this level of detail before in Spain, an area with diverse forests and where water management for ecosystems and society is a pressing issue. The text and figures are generally clear, although I provide some suggestions to improve clarity in my specific comments below.

Thank you for your positive comments.

Q2.1. My major criticism of the paper is that it could have gone beyond a modelling exercise and aim at gathering more evidence that could have provided (i) a more robust spatial evaluation of the model or (ii) additional support to the paper's conclusions. I'm aware that there are no eddy flux or sap flux stations with a sufficient geographic coverage to be representative of the entire country, precluding an extensive model evaluation. However, I wonder whether products such as FLUXCOM ET, with a relatively good spatial resolution (Jung et al. 2019), could be used to address some of the questions in the paper. This is a gridded product which I believe would be relatively straightforward to use in the paper and it also provides estimates of GPP, which aid in the interpretation of some of the results (see my specific comments below). I think that showing results from the modelling alongside FLUXCOM would make this a more valuable contribution.

R2.1. Thank you for your suggestion. We agree that a lack of comparison/evaluation section was weakening our contribution. We conducted a comparison between the MEDFATE results and the evapotranspiration from GLEAM and the blue water from SIMPAL in a new section (2.7.). GLEAM and FLUXCOM have a similar pattern and the correlation is very high according to Jung et al., 2019.

"2.7 Model evaluation

MEDFATE predictions have already been evaluated at the forest stand scale in terms of soil moisture dynamics, plant transpiration and water status in Mediterranean forest (De Cáceres et al., 2021, 2015). Given the focus of the present work, we evaluated regional-scale patterns of green and blue water predicted by MEDFATE against those produced by alternative methods. First, we did a comparison between the results for average blue water of MEDFATE with the average blue water from the Precipitation Runoff Integrated Model (SIMPAL acronym in Spanish) of the Spanish government. SIMPAL model is calibrated with stations that measure stream flows across Spain and it is interpolate a 1 km^2 resolution for the whole country (Estrela et al., 2012). Second, we compared green water patterns predicted by MEDFATE against those of GLEAM, which is derived from satellite data and covers the world with a resolution of 0.25 degrees (Martens et al., 2017). We used the GLEAM v3.7a that defines the evapotranspiration as the sum of transpiration (from short and tall vegetation), interception (from tall vegetation), soil and open water evaporation and snow sublimation. The

comparative indicated that the blue water of the SIMPAL model and the green water of GLEAM followed the same regional patterns that MEDFATE (see Figure S7 and S8). SIMPAL reaches higher values of blue water in the Atlantic biome and lower values in west Temperate Mediterranean than MEDFATE. In opposition GLEAM evapotranspiration is higher in the west Temperate Mediterranean and lower in the Atlantic forest than MEDFATE. MEDFATE models at stand scale whereas SIMPAL and GLEAM models are at regional scale. At stand scale the evapotranspiration is high in the Atlantic forest and then the blue water is lower. West Temperate Mediterranean is characterized by forest of low basal area and LAI (Table S2, Figure S4). Therefore the evapotranspiration is lower and blue water higher than surrounding forest. In other words, the differences observed seem to arise from the fact that MEDFATE takes into account the stand structural characteristics, which are difficult to represent in models based on interpolation (SIMPAL) or remote sensing (GLEAM).”

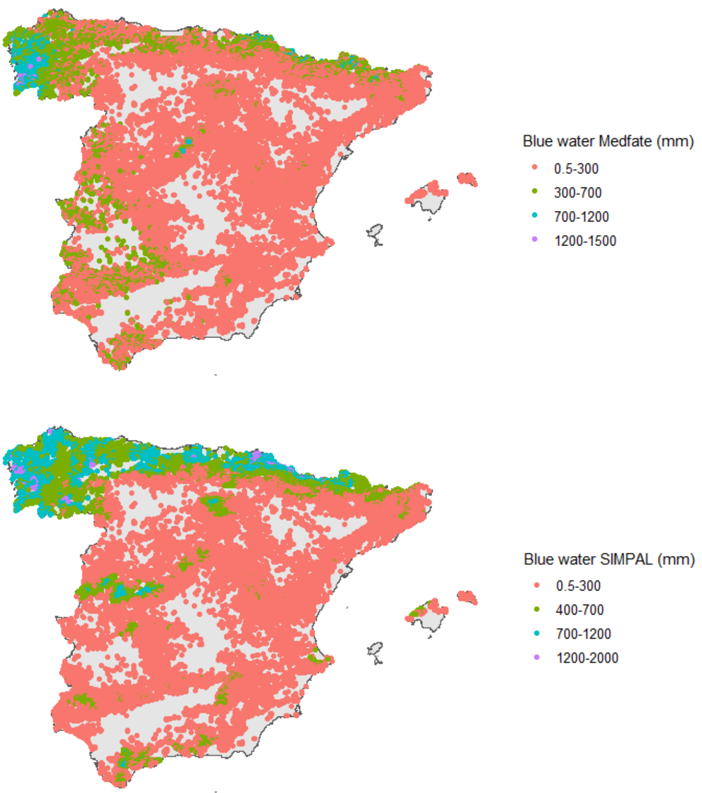

Figure S7. Average blue water maps for SFI3 plots according to MEDFATE and SIMPAL (Spearman correlation=0.70, R^2=0.354).

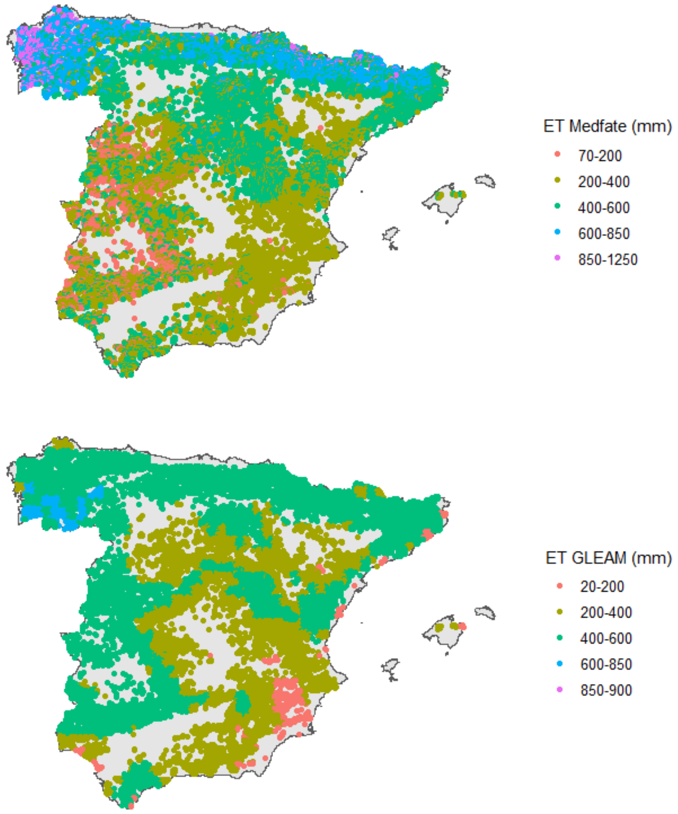

Figure S8. Average evapotranspiration (ET) maps for SFI3 plots according to MEDFATE and GLEAM (Spearman correlation=0.62, R^2=0.218)

Figure S4 is showed in R2.14.

Q2.2. If the authors decide to stick to the model results, then I think that the title of the paper should explicitly say that it's a modelling exercise. For example 'Regional patterns and drivers of modelled water flows along environmental, functional and stand structure gradients in Spanish forests'. Or maybe mention the combination of modelling and forest inventory data, which, to me, is also a nice contribution of the paper and that would remain hidden if only 'model' appears in the title. Anyway, these are suggestions that put the focus more on the methodology and tools, which seems to align with what the paper currently shows. If, alternatively, the focus is indeed on the actual results and geographical patterns, the additional suggested analyses above would make this contribution more valuable, in my opinion.

R2.2. We changed the title to: "Regional patterns and drivers of modelled water flows along environmental, functional and stand structure gradients in Spanish forests". We think that this title shows the content of the article.

Specific comments

Q2.3. The abstract reads a bit vague, I think it that someone skimming the abstract would expect some overall quantification of green and blue water for different forest types and different seasons for example. Lines 12-19 simply describe methods and it feels a bit excessive, so you could make these descriptions a bit more compact and free some space to provide more quantitative results. Also, the last sentence of the abstract, especially the last part is not very clear to me. Why do you say LAI is a filter for excess water? Do you mean that LAI influences the partitioning of blue and green water, for example?

R2.3. We have removed or gathered some sentences of the methodology and we have included the relative blue water percentages of biomes and forest functional groups. We changed the conclusion: "This study highlights how the green water is decoupled of the blue water across the forests typologies of whole Spain and it how the species functional traits (deciduous vs evergreen) can influence the blue water production."

Q2.4. L. 49. The Jasechko study was a bit controversial (Coender-Gerrits et al. 2014), I would suggest you use other references to support the dominant role of transpiration in terrestrial ET for specific biomes (Schlesinger & Jasechko 2014; Wei et al 2017). Also, please be sure to provide the correct numbers for the estimations of green water when comparing with estimations that only provide transpiration or T/ET. I understand that green water also includes evaporation from vegetated and other surfaces.

R2.4. In the sentence (L52/53) we talked about the proportion and dominant role of the green water (ET) about the total precipitation but not the transpiration in the terrestrial ET. But we changed the reference to Schlesinger & Jasechko 2014. The ratio between green water and precipitation that we cited (40-70 %) is correct according to Schlesinger & Jasechko, 2014.

Q2.5. L. 61 - 62. I think that the focus here is more on 'water saver' vs 'water spender' species, which is not always related to isohydric-anisohydric dimension (Martínez-Vilalta & Garcia-Forner 2017). I would simply mention the distinction between water spender/water savers or loose vs strict stomatal regulation of transpiration.

R2.5. Thank you for the suggestion. We agree that isohydric-anisohydric dimension can be inadequate. We changed the reference to water saver/spender in the sentence.

Q2.6. L. 72. When reading here 'forest species traits' one would expect that this is dealt with in this paper, but this is not the case. It's more a functional type approach.

R2.6. We changed "forest species traits" by "functional forest groups".

Q2.7. L.133. Is it E-OBS or Worldclim data you are using? As I understand, E-OBS contains already the daily data you need, and the link is not the same as the official E-OBS one (https://www.ecad.eu/download/ensembles/download.php); could you clarify this?

R2.7. We used E-OBS daily data that was downscaled with the Worldclim resolution (more details in Moreno and Hasenauer, 2016). The link that we included is the correct one. We removed the reference to Worldclim to avoid confusion.

Q2.8. L. 136. Provide website URL for IGN.

R2.8. We included the URL.

Q2.9. L. 138 - 141. I would use 'climatic moisture', to differentiate from 'soil moisture', for example. And for 'seasonality' use 'precipitation seasonality', also in figures and other instance throughout the text (abbreviated, if needed).

R2.9. We changed the two names along the text.

Q2.10. L. 173. These are aTmax and bTmax in De Cáceres et al 2023, right? As it's written, it's not very clear how maximum transpiration is estimated. Given the relevance of LAI in these calculations (which impact on the results you observe), maybe it's worth including the equations relating max transpiration and LAI here, and explain clearly how species-specific parameters are obtained.

R2.10. We included more information about how maximum transpiration was estimated: "The estimation of maximum transpiration for the entire stand (Emax,stand), excluding considerations for soil water deficit, relies on the daily Penman's potential evapotranspiration (PET) and an empirical relationship established by Granier et al. (1999). But MEDFATE modified the Granier equation with aTmax and bTmax: If it is assumed that the entire leaf area of a stand corresponds to a single cohort i, the equation is:

$$\frac{E_{\max,\text{stand}(i)}}{\text{PET}} = a_{\text{Tmax}} \cdot \text{LAI} + b_{\text{Tmax}} \cdot \text{LAI}^2,$$

where aTmax and bTmax represent species-specific parameters for cohort i. Assuming reliable species-specific estimates are accessible for aTmax and bTmax, the equation can be applied to calculate Emax,stand(i), which denotes the maximum stand transpiration when dominated by the species of cohort i. Once Emax,stand(i) is determined for each species in the stand, the portion of SWR absorbed by a particular cohort i is employed to estimate its maximum transpiration (Emax(i)) from Emax,stand(i) (Korol et al., 1995)."

Q2.11. L. 191-193. Please provide the source of these allometries, if published, or provide them in a supplementary table if you have derived them for this study. If the latter, please also explain briefly which data sources were employed and the overall data design (number of tree/plot replicates, geographical scope – Spain or also data from elsewhere).

R2.11. The details are explained in the Supplementary material of De Cáceres et al., 2023. We included the reference to the tables of Supplementary material of this work (De Cáceres et al., 2023) where you can find the information of the source and values of the allometries: "Taxon-specific parameter details are shown in supplemental material of De Cáceres et al., 2023."

Q2.12. L. 209.'functional group'

R2.12. Done.

Q2.13. L. 229. How was this R2 estimated?

R2.13. It is only: 1-(Sres/Stot), being Sres the sum of the squared differences between the observed values and the predicted values, and Stot is the total sum of squares, which is the sum of the squared differences between the observed values and the mean of the observed values.

Q2.14. L. 241. Readers not familiar with Spain's geography will not know where these ranges are. I would suggest providing a more complete figure A1, with several informative layers, as well as the current climatic biomes. These would need to include topography, distribution of forest types and also the important spatial predictors in your models from Fig 5 (i.e. multiple small maps in fig A1).

R2.14. Thank you for the suggestion. We included more maps with the altitude, functional forest group and the three main predictors of the models. Theses maps were included in the Supplemental Material since they occupied a lot of space:

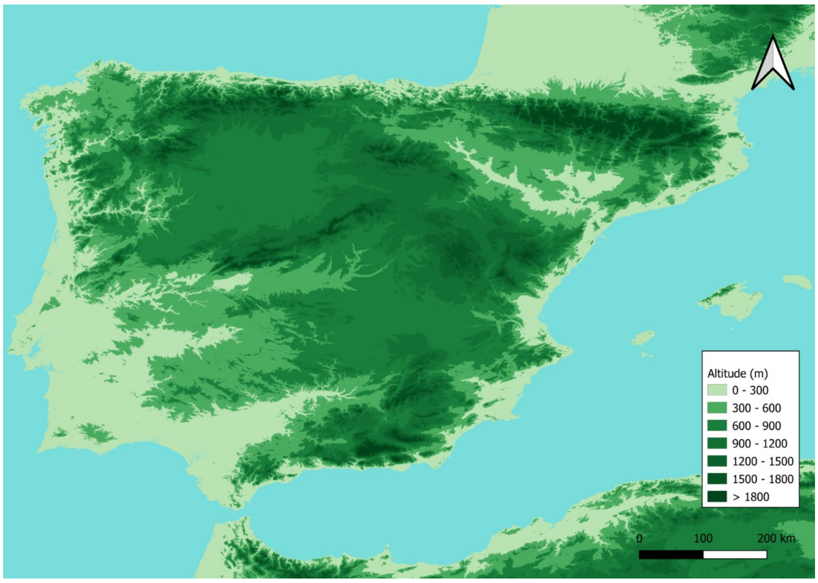

Figure S2. Altitude in the Iberian Peninsula and Balearic Islands.

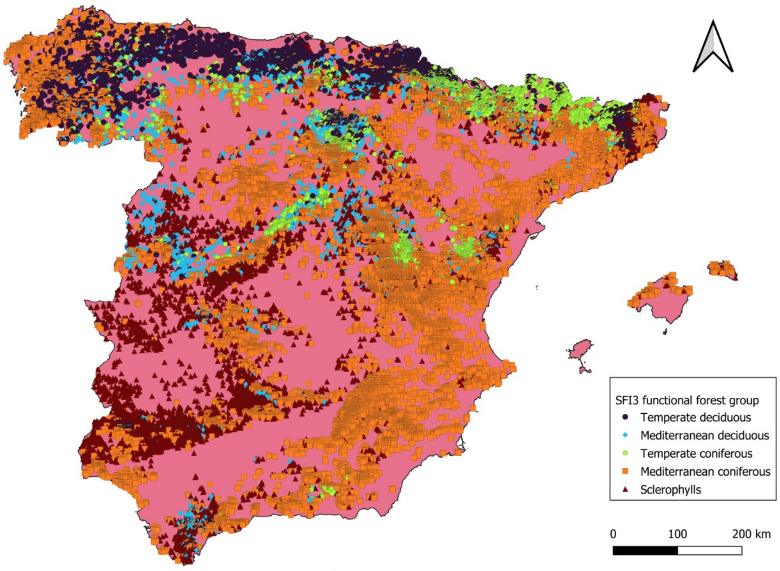

Figure S3. SFI3 plots by functional forest group

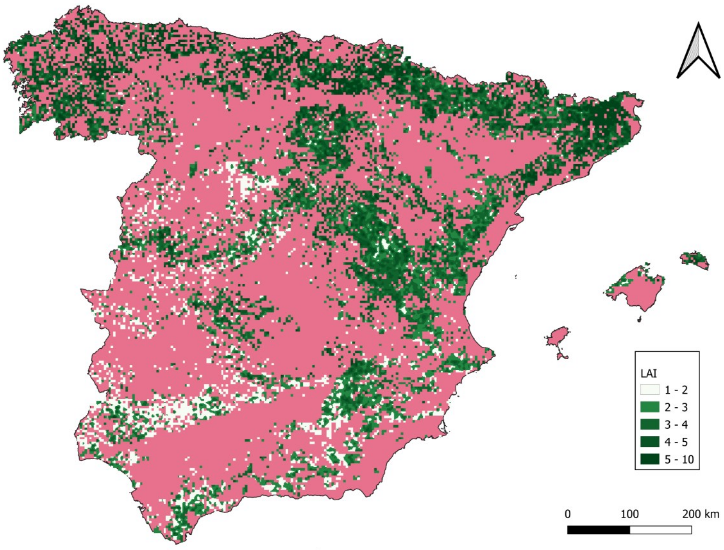

Figure S4. LAI of the Spanish Peninsula and Balearic Islands forest.

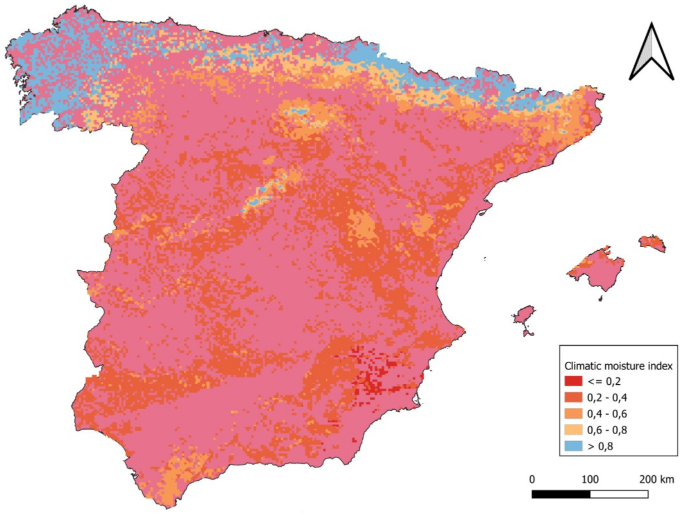

Figure S5. Climatic moisture index in Spanish Peninsula and Balearic Islands.

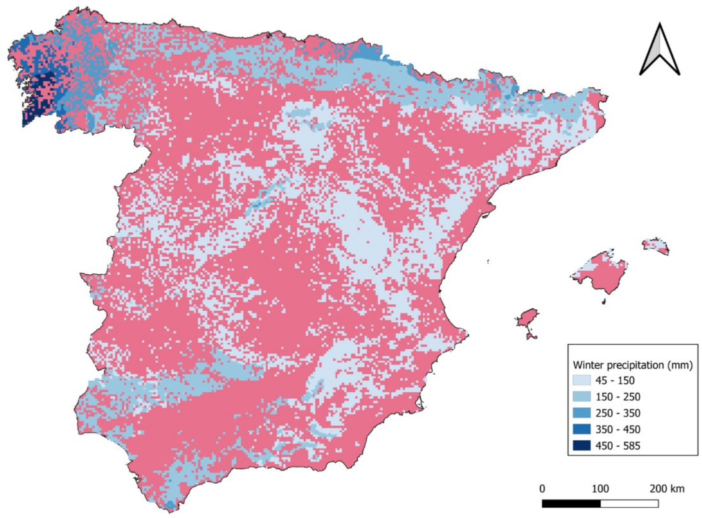

Figure S6. Winter precipitation in Spanish Peninsula and Balearic Islands.

Q2.15. Figure 2 and Figure 3. I think it would be informative to show in these figures also the % of soil evaporation and interception. Even Figure 4 could show the contributions of T vs interception + soil evaporation (these two could be combined for clarity) by showing this within the 'blue water' bar.

R2.15. We included in the Supplemental Material the % of interception and soil evaporation related to precipitation by biomes and functional groups (Figure S9-13). We did not include the interception/soil evaporation in Figure 4 because of the image is very full. But with the new Supplemental Material we show the ratio of interception and soil evaporation in the green water.

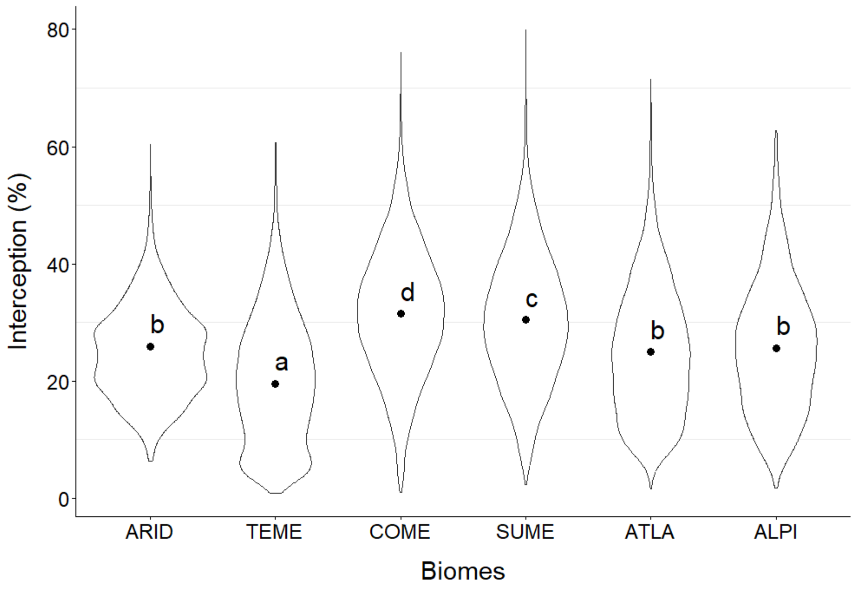

Figure S9. Percentage of interception in the different biomes. Different letters denote significant differences among biomes ($p < 0.01$) after Tukey's test. Abbreviations of the biomes: ARID, Arid; TEME, Temperate Mediterranean; COME, Continental Mediterranean; SUME, Submediterranean; ATLA, Atlantic; ALPI, Alpine.

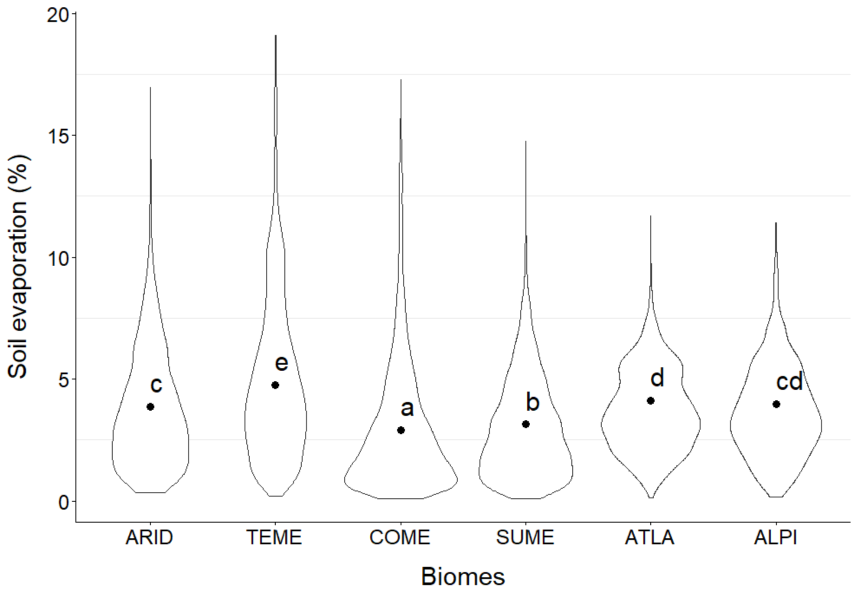

Figure S10. Percentage of soil evaporation in the different biomes. Different letters denote significant differences among biomes ($p < 0.01$) after Tukey's test. Abbreviations of the biomes: ARID, Arid; TEME, Temperate Mediterranean; COME, Continental Mediterranean; SUME, Submediterranean; ATLA, Atlantic; ALPI, Alpine.

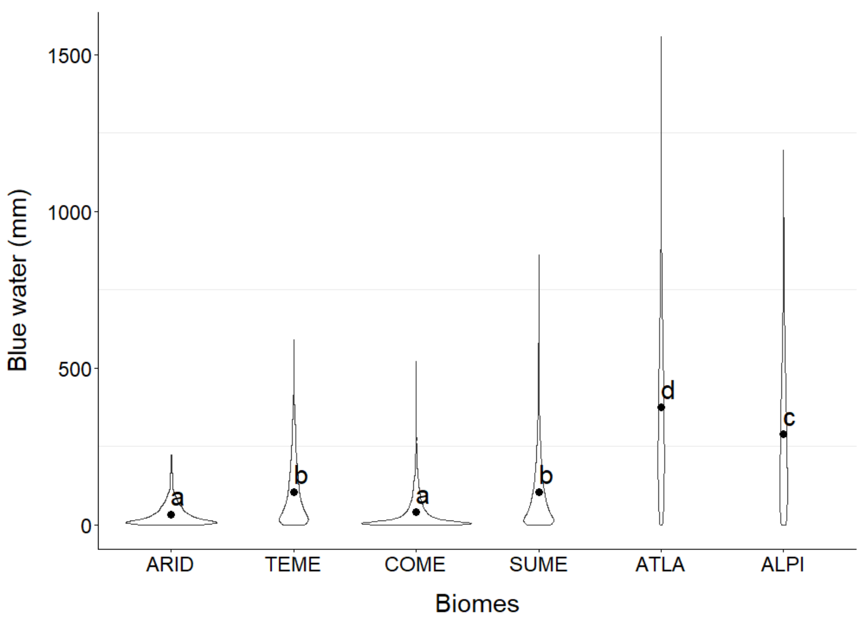

Figure S11. Blue water amount in the different biomes. Different letters denote significant differences among biomes ($p < 0.01$) after Tukey's test. Abbreviations of the biomes: ARID, Arid; TEME, Temperate Mediterranean; COME, Continental Mediterranean; SUME, Submediterranean; ATLA, Atlantic; ALPI, Alpine.

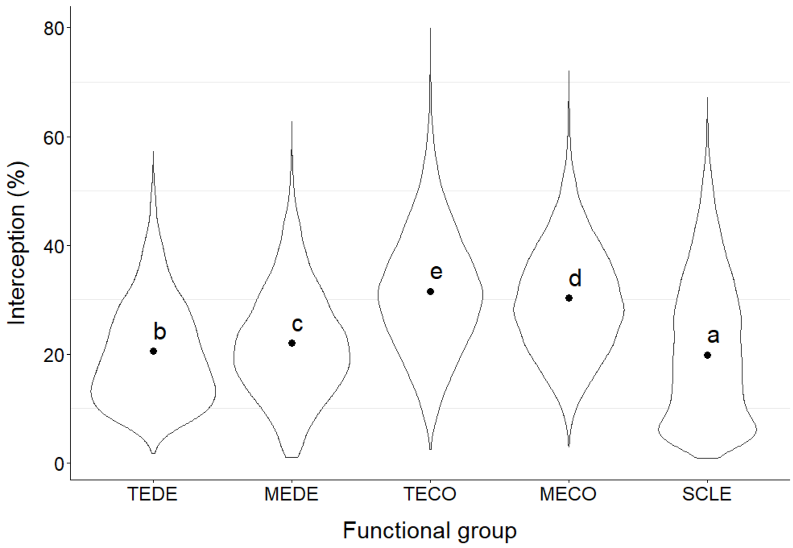

Figure S12. Percentage of interception of the different functional groups of species. Different letters denote significant differences among forest types ($p < 0.01$) after Tukey's test. Abbreviations of the functional groups: TEDE, Temperate deciduous; MEDE, Mediterranean deciduous; TECO, Temperate coniferous; MECO, Mediterranean coniferous; SCLE, sclerophylls.

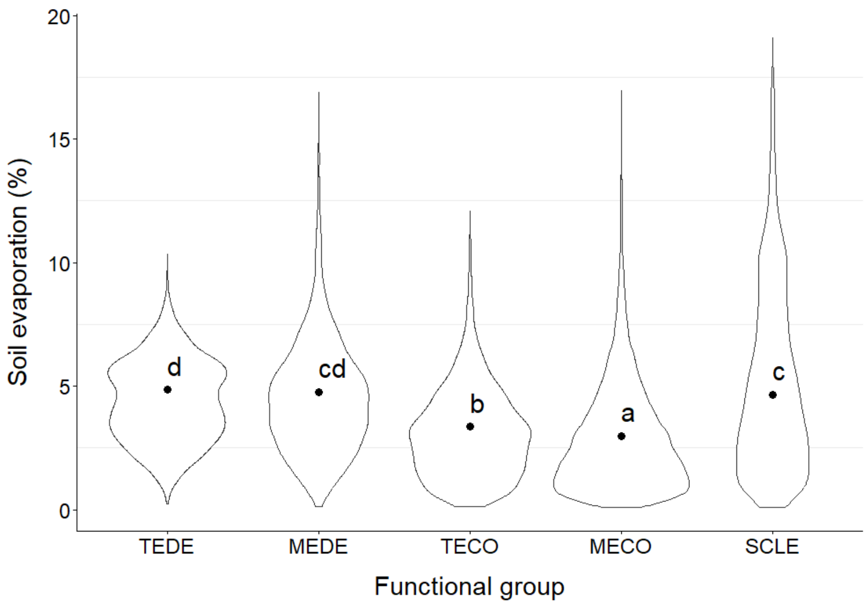

Figure S13. Percentage of soil evaporation of the different functional groups of species. Different letters denote significant differences among forest types ($p < 0.01$) after Tukey's test. Abbreviations of the functional groups: TEDE, Temperate deciduous; MEDE, Mediterranean deciduous; TECO, Temperate coniferous; MECO, Mediterranean coniferous; SCLE, sclerophylls.

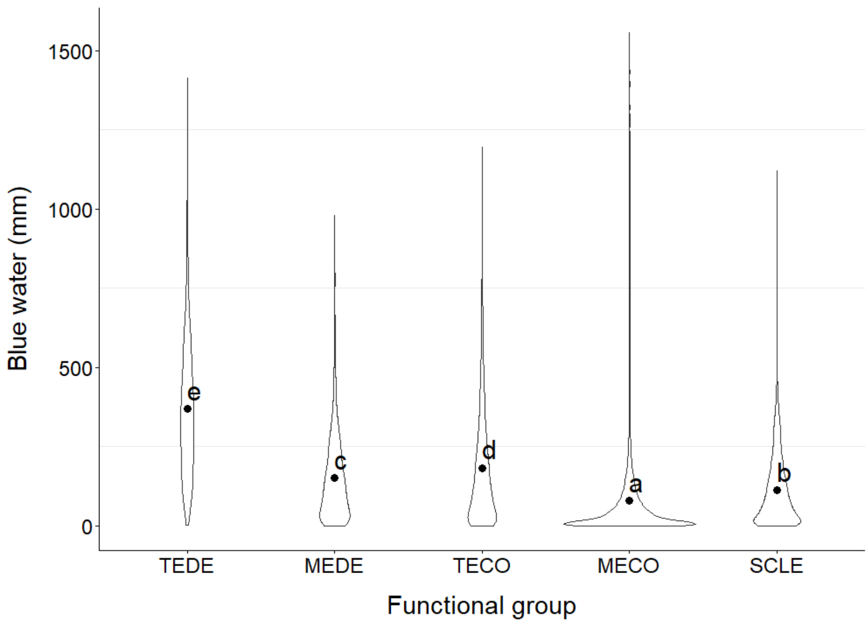

Figure S14. Blue water amount of the different functional groups of species. Different letters denote significant differences among forest types ($p < 0.01$) after Tukey's test. Abbreviations of the functional groups: TEDE, Temperate deciduous; MEDE, Mediterranean deciduous; TECO, Temperate coniferous; MECO, Mediterranean coniferous; SCLE, sclerophylls.

Q2.16. Related to this, I also wonder whether you could check whether the results agree with the global patterns observed for the partitioning of terrestrial evaporative fluxes: a plot of interception, transpiration and soil evaporation, as a fraction of rainfall following Good et al. 2017 paper (Fig. 1), as a function of climatic aridity.

R2.16. We tried to replicate the figure and we included the new information in the section 4.1 and Supplemental Material: "Transpiration is the main component of the green water in arid environments and decrease in wetter regions where the interception increases their importance (see Figure S15). This result is concordant with the global patterns of green water flux where the transpiration is bigger in arid environments (Good et al., 2017)."

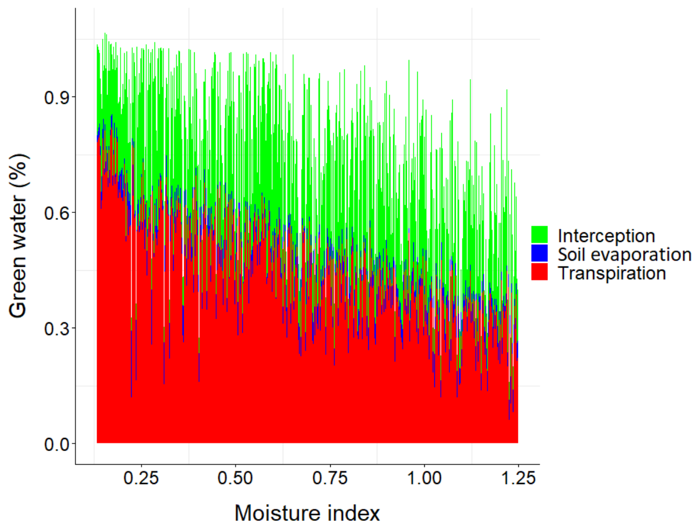

Figure S15. Green water partitioning (interception, soil evaporation and transpiration) along the moisture index gradient.

Q2.17. L. 350.'water evapotranspired or intercepted by the canopy'

R2.17. Done.

Q2.18. L. 365-366. The model provides the ET partitioning so you could show T,soil evaporation and interception to support this interpretation, I believe.

R2.18. In the Figure S15 previously referenced we showed the partitioning of the ET in their three components.

Q2.19. L. 382. This statement on the anisohydry is too general to be used here and I would be more cautious; moreover, I don't think this pattern emerges from Klein's paper. Deciduous species such as Populus or Quercus robur (Martínez-Vilalta et al. 2014; Urli et al 2014) can hardly be considered isohydric, and some evergreen species within *Juniperus* can be quite anisohydric.

R2.19. In the text we said that the deciduous are mainly anisohydric. But it is true that there are many exceptions and deciduous species can isohydric and evergreen can be anisohydric. We removed the reference to aniso/isohydric to avoid confusion.

Q2.20. L. 404. 'produce more blue water'

R2.20. Done.

Q2.21. L. 428. ' a poor predictor'

R2.21. Done.

Q2.22. L. 441 - 443. But in fact this detailed partitioning of the fluxes has not been addressed in the paper.

R2.22. We considered that gathering the fluxes in green/blue water it was the best way of realized the analyses and the work. But we included more information in the updated manuscript in the Supplemental material about the green water components and their differences by biomes, functional group or along the moisture index gradient (Figures S9-15 previously showed).

Q2.23. L. 450-460. I don't find this part of the conclusions particularly meaningful. In fact,most of the conclusions before this section revolve around the limitations of your approach (a model, not observations) and I think you could highlight better what are the broader implications of your exercise, both in methodological terms and for the results you obtain. What is the power of combining actual forest structure data with a process based model to estimate forest water fluxes in time? i.e. incorporating variation in forest structure through repeated NFI or other airborne/remote sensing surveys, for example.

R2.23. We included more information in the conclusion about the inclusion of forest structure data: "In this work we showed that stand structure variables and functional composition play a substantial role in ecohydrological simulations. Moreover, the incorporation of forest

inventories to these models allow to do analyses at large spatial scales across multiple climates, forest typologies and species composition."

References:

De Cáceres, M., Martínez-Vilalta, J., Coll, L., Llorens, P., Casals, P., Poyatos, R., Pausas, J. G., and Brotons, L.: Coupling a water balance model with forest inventory data to predict drought stress: the role of forest structural changes vs. climate changes, Agricultural and Forest Meteorology, 213, 77–90, https://doi.org/10.1016/j.agrformet.2015.06.012, 2015.

De Cáceres, M., Mencuccini, M., Martin-StPaul, N., Limousin, J.-M., Coll, L., Poyatos, R., Cabon, A., Granda, V., Forner, A., Valladares, F., and Martínez-Vilalta, J.: Unravelling the effect of species mixing on water use and drought stress in Mediterranean forests: A modelling approach, Agricultural and Forest Meteorology, 296, 108233, https://doi.org/10.1016/j.agrformet.2020.108233, 2021.

Estrela, T., Pérez-Martin, M. A., and Vargas, E.: Impacts of climate change on water resources in Spain, Hydrological Sciences Journal, 57, 1154–1167, https://doi.org/10.1080/02626667.2012.702213, 2012.

Good, S. P., Moore, G. W., and Miralles, D. G.: A mesic maximum in biological water use demarcates biome sensitivity to aridity shifts, Nat Ecol Evol, 1, 1883–1888, https://doi.org/10.1038/s41559-017-0371-8, 2017.

Granier, A., Bréda, N., Biron, P., and Villette, S.: A lumped water balance model to evaluate duration and intensity of drought constraints in forest stands, Ecological Modelling, 116, 269–283, https://doi.org/10.1016/S0304-3800(98)00205-1, 1999.

Jung, M., Koirala, S., Weber, U., Ichii, K., Gans, F., Camps-Valls, G., Papale, D., Schwalm, C., Tramontana, G. and Reichstein, M., 2019. The FLUXCOM ensemble of global land-atmosphere energy fluxes. *Scientific data*, *6*(1), p.74.

Korol, R. L., Running, S. W., and Milner, K. S.: Incorporating intertree competition into an ecosystem model, Can. J. For. Res., 25, 413–424, https://doi.org/10.1139/x95-046, 1995.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geoscientific Model Development, 10, 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

Moreno, A. and Hasenauer, H.: Spatial downscaling of European climate data, International Journal of Climatology, 36, 1444–1458, https://doi.org/10.1002/joc.4436, 2016.