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

RC1: ‘Comment on hess-2023-255’, Anonymous Referee #1, 05-12-2023

General comment:

This work can represent a very significant contribution to improve our knowledge about the effect of forest structure and climate on water resources (or blue and green waters) in a very heterogenous context of climates and forest typologies such as Spain. Although it is true that it has been proposed at country level, in my opinion, the results presented are those expected in terms of the differences observed between the biomes and the types of forests present in the study area, which gives credibility to the results from modelling apart from other about the validation of results based on field data (transpiration, water in the soil, runoff, etc.). Its subsequent use in studies at higher spatial resolution (at the basin scale, for example) could help managers to better define forest management with a strong hydrological foundation (to improve green water, blue water, or both) and therefore promoting ecohydrology-based management to improve water resource availability at regional scales.

I value very positively the great effort made in terms of obtaining information for the characterization of the forest inventory plots, as well as other information necessary to apply the model on a scale as ambitious as Iberian Spain, with such diverse forest typologies and climates and extension.

Thank you very much for your positive comments.

Q1.1. However, I believe that the work still needs to be greatly improved to be finally published in a journal like HESS. In this sense, I consider very important to justify the model performance to produce realistic estimates for all the forest typologies addressed in the work. Although it is true that a recent work is indicated (De Cáceres et al., 2023) where the validation for some species with experimental field data is presented, this work should at least indicate how the latter has addressed the calibration and validation of the model, and not only relegating this information to the conclusions section. As I indicated, I understand this is not straightforward given you have not experimental plots in all the forest typologies, but I understand more efforts should be carried out to justify the use of the model for the entire territory of Spain and the validity of their results. In this sense, maybe you can consider available databases to discuss about the model performance (such as that for sapflow from https://essd.copernicus.org/articles/13/2607/2021/) when comparing your estimates to some hydrological component such as transpiration or runoff.

R1.1. We agree that a validation in more tree species or biomes can be necessary since De Cáceres et al. (2015) and De Cáceres et al. (2021) didn’t cover the whole country and all the forest typologies and the evaluation of De Cáceres et al. (2023) focused on forest dynamics. Nevertheless, there are a lack of field work data at

country scale. Our current efforts include evaluating the model in other (temperate) forest types, using datasets like PROFOUND. However, we felt that for this paper the model evaluation should be made at larger scales, in accordance with the scale of application. Thus, we have included a comparison with two flow water products, for green water and blue water respectively, at regional scale that can be more adequate for the scale of this study. We answered with more details in the R1.13.

On the other hand, although the fundamentals of the model are understood quite clearly, I believe that the work should not simplify relevant information to understand how the calculations have been carried out in the study, either in the materials and methods section or as additional information. For example, it is not explained how the LAI is calculated or how the data obtained for the IF3 plots are regionalized or how the modeled fluxes are extrapolated to the entire Spanish forest territory (except for the Canary Islands). In my opinion, this aspect must be clearly addressed, since in previous works related to your model performance, I am not able to understand how this is addressed. I believe this work is the correct place to give a proper explanation about how spatial dimension is considered within your model framework.

R1.2. We answered the question about the calculation of the LAI in the R.1.23. The question about the regionalization was answered in R1.15.

I believe English flows well and the manuscript is easy to be followed due to its clarity and well structure. I have, however, several suggestions and comments which have been added as comments within the manuscript. In this sense, apart from considerations about the modelling approach, I would recommend reviewing the introduction section for improving it (see please my comments on it).

Q1.3. L22: I recommend including values for this LAI plateau depending on forest type/climate. Take into account you did not include any quantitate information through the entire abstract, thus I would add more "numbers" to improve it in order to be less qualitative

R1.3. Thanks for the suggestion. We have included the values in the abstract: (around 2.5-3). The plateau can be seen for each functional group in Figure 6. Additionally, we have added some more quantitative results in the abstract.

Q1.4. L25: Too general conclusion, Please be more specific based on your results

R1.4. We changed the conclusions to: "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."

Q1.3. L34: Remove "On the other hand"

R1.3. Done.

Q1.4. L36: please include reference to green water

R1.4. Done: (Llorens et al., 2011)

Q1.5. L43: please define LAI; also, it is not the only variable used for characterizing stand structure. You need further justify why LAI can be the most optimum variable to describe stand structure rather than density or cover, among others

R1.5. We have included the definition of LAI in the introduction: "Stand forest structure is also key in rainfall partitioning through its relationship with stand LAI (total area of leaves of the canopy per unit horizontal ground area) which is a key variable determining transpiration (Granier et al., 2000) and/or the interception of vegetation." In the Methods section (Statistical analysis) we have explained why LAI (and also BA) are the variables used to describe stand structure: "Basal area serves as an indicator of canopy cover, as it is calculated by summing the diameters at breast height (dbh) of all trees per hectare, and higher basal area values can result in increased interception. LAI determine the transpiration of the trees through of the leaf surface."

Q1.6. L50: I miss more recent works in Mediterranean forests, and specially those carried out in Spain

R1.6. We included one field work carried out in a Mediterranean forest in Spain and another one in a Mediterranean forest in Israel: (Campos et al., 2016; Qubaja et al., 2020)

Q1.7. L67: I don´t understand the logic followed here; you first focused on two really contrasted scales and then you said that the effects of stand structure and traits are not well examined in the literature.

R1.7. We removed the sentences about the stand structure and traits to avoid the confusion.

Q1.8. L74: In my opinion, this is the key message of your introduction, However, I consider you should further review more recent articles in order to better justify your objectives but also improving the statements of your introduction. I find the introduction two general and not really supported by recent literature.

R1.8. Thanks for your suggestion. We added more information in the text: "The stand structure variables have been studied at local scale (Benyon et al., 2017; Simonin et al., 2007), but the landscape ecohydrological simulations have been carried out without incorporating the role of the stand structure variables or the differences in species and functional composition (Hoek van Dijke et al., 2022; Mastrotheodoros et al., 2020)." We hope the statement of our objectives is clearer now.

Q1.9. L95: You should also include information regarding geology, soils, geomorfology, Variation in climatic aspects is important but also other aspects should be added to give a correct general description of your study area.

R1.9. Thank you for your comment. But we did not include information about geology or soils because the basic descriptions of geology/soils in Spain mainly focus on the difference between acidic and basic pH, which is not relevant for this work.

Q1.10. L101: please include references to these climatic values

R1.10. The climatic values were derived of the average climatic extracted for the SFI3 plots. But we removed the values to avoid confusion.

Q1.11. L111: LAI is in the end the most important stand parameter you considered. In this sense, you should include a detailed explanation about how it is estimated based on SFI data or whatever you considered.

R1.11. Thanks for pointing at this shortcoming of the original text. The variables and calculations that are based on SFI data are described in detail in a new section:

"2.5. Parameter estimation

Data from forest inventory plots included tree height (*H*) and tree diameter at breast height, which was used to obtain estimates of foliar biomass (hence leaf area after multiplying by *SLA*) and crown ratio (*CR*) via species-specific allometries (see Table S1-3 of De Cáceres et al., 2023 for more details). In the model, *SLA* (Specific Leaf Area (sq *mm/mg*)), *the ratio of leaf area to leaf dry mass,* is constant for every species (De Cáceres et al., 2023). Leaf area index (LAI in $m^2 \cdot m^{-2}$) was calculated from the foliar biomass (in $kg \cdot m^{-2}$) by using a *specific leaf area* coefficient (*SLA*, in $m^2 \cdot kg^{-1}$) that is species-specific (LAI = foliar biomass * *SLA*). Taxon-specific parameter details are shown in supplemental material of De Cáceres et al., 2023. Soil data of each forest inventory plot was extracted from the SoilGrids database (Hengl et al., 2017). For all plots four soil layers down to a total depth of 4 m were initially considered, but the deepest layers were merged into a rocky layer (95 % of rocks) following the depth of the R horizon. A monotonous increase in rock fragment content across soil layers from the surface to the rocky layer was defined based on surface stoniness classes determined in SFI3 plot surveys."

Q1.12. L144: nothing is said about blue and green water calculations. I understand runoff and soil percolation are considered as blue water but further information is required to understand how this is addressed on your study, In addition, what about soil evaporation as key component of green water? I would suggest to include a final paragraph explaining this

R1.12. We have included an explanation of the calculation of blue and green water in a new section:

"2.6. MEDFATE simulations:

MEDFATE was run on each selected SFI3 plot using daily weather data (temperature and precipitation, PET, radiation and relative humidity) corresponding to a 10-yr period centered on the year of the SFI3 sampling (1997-2008). We calculated the blue water as the sum of the runoff and deep water and the green water as the evapotranspiration (that included the sum of the transpiration, interception and soil evaporation). The percentage of interception and soil evaporation about green water now are showed in the Supplemental material (Figure S7-S12)"

Q1.13. L154: I think further work should be presented here in order to justify the use of this model to simulate all the forest typologies presented in Spain, since nothing is said about calibration and validation ot the model in the manuscript. I understand the model has been properly calibrated and validation for some of the forests but not all the studied ones within the manuscript. For those, you should clearly justify how your model estimates can be considered close to reality

R1.13. We cannot validate in all Spain due to the lack of field data for them. But we included a paragraph in a new section (2.7) where we did a comparative between the results of MEDFATE with the blue water from the ministry and the evapotranspiration from GLEAM. The map comparatives and statistical correlations are showed in the Supplemental material (Figure S4 and Figure S5). Theses maps showed the same regional patterns for these sources with MEDFATE. Although we know that a validation with field data it is better at country scale it is very harsh.

Q1.14. L186: how did you incorporate these physiological parameters within your study? How did you calculate stand transpiration based on parameters which are depending on tree species?

R1.14. We included more information about how stand 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:

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

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).”

Q1.15. L188: I believe you should explain how your model is extrapolated to the whole Spain mainland based on SFI3 data. see in the figure 1 you finally considered a square pixel as your spatial domain, but noting is said about this in your M and M section. I understand you run your model for each IF polygon and then you calculate each water cycle component at a daily scale. However, from this water balance estimate to whole territory, there are several steps you should accomplished and they are not clearly explained in your study. Please consider to improve the explanation when moving stand scale into landscape scale for obtaining blue water variables.

R1.15. We ran the MEDFATE model for every SFI3 plot. The results are at daily scale and we averaged the values (365 days x 10 years) for the analyses. In the maps we showed theses averages values after we rasterized the SFI3 plots. In the rasterization we transformed a vectorial point database (the SFI3 plots) in a raster pixel of 1 km^2. We did not an extrapolation since every pixel have the same value of the SFI3 plot where it was transformed. The SFI is distributed with a systematic survey along the forested areas with a density of ~1 plot/km^2. Then we don’t need extrapolate the data since it cover the main forest areas of Spain. The pixel without a SFI3 plot don’t have values (it is the reason because the map do not cover the whole country since many pixels are “NA” where there are not forest areas). The rasterization is only a visual way of show the same values of the vectorial SFI3 plot database. In the Figure S5 (referenced in R1.13) we show the results of the blue water comparison between MEDFATE with SIMPAL model. In this case we did not transform the vectorial data in a raster pixel. But the data are the same that in the map of Figure 1A. Therefore, the rasterization it was only an aesthetic decision. However we included the information about the rasterization in the Figure 1: “Maps were realized with a rasterization of the SFI3 plots results at 1 km^2.”

Q1.16. L208: include topography as well

R1.16. Done.

Q1.17. L234: I would consider to modify this title to be clearly different than the next one. Something like this "Spatial patterns of blue water of Spanish biomes" could be applicable

R1.17. Thanks for your suggestion. We have changed the title to: “Spatial patterns of blue water in the Spanish biomes”

Q1.18. L238: Why did your consider this value as your threshold? It is based on a reference, on your data distribution? Please clarify this

R1.18. It was a descriptive value according to the maps. We changed the sentence to: “had values between 300-500 mm”

Q1.19. L241: I would suggest to interpret your blue water estimates based on mm and % of precipitation in two separated parts. These metrics are indicating different aspects, and thus I think they should be individually treated

R1.19. We included more explanations about the blue water amount (mm) but the patterns are very similar with blue water percentages.

Q1.20. L251: what about absolute values? are they indicating the same or not? I think that including another table for water depth is very pertinent.

R1.20. We included the Tukey test for the absolute values in the Supplemental material (Figure S9 and Figure S12). The results are very similar to the relative blue water.

Q1.21. Table 1. further explanation about differences on estimates values are required. For example, which is the meaning of a value of -1.9 instead of 0.22? I would include more information about beta regression coefficients in the statistical analyses section

R1.21. We included the explanation in the Table 1 and Table 2: “A positive value indicates that an increment occurred in the biome with respect to the observed values in the intercept”.

Q1.22. Table A2. you should include information regarding the three last columns

R1.22. The information about the indices was explained in the section 2.3. The indices do not have units.

Q1.23. Figure 2A: how was LAI simulated? This aspect should further explained, and especially, given the high importance of this parameter when explaining blue water differences among forest types

R1.23. We have included this explanation: “Leaf area index (LAI in $m^2 \cdot m^{-2}$) was calculated from the foliar biomass (in $kg \cdot m^{-2}$) by using a *specific leaf area* coefficient (SLA, in $m^2 \cdot kg^{-1}$) that is species-specific (LAI = foliar biomass * SLA).”

References:

Benyon, R. G., Nolan, R. H., Hawthorn, S. N. D., and Lane, P. N. J.: Stand-level variation in evapotranspiration in non-water-limited eucalypt forests, Journal of Hydrology, 551, 233–244, https://doi.org/10.1016/j.jhydrol.2017.06.002, 2017.

Campos, I., González-Piqueras, J., Carrara, A., Villodre, J., and Calera, A.: Estimation of total available water in the soil layer by integrating actual evapotranspiration data in a remote sensing-driven soil water balance, Journal of Hydrology, 534, 427–439, https://doi.org/10.1016/j.jhydrol.2016.01.023, 2016.
De Cáceres, M., Molowny-Horas, R., Cabon, A., Martínez-Vilalta, J., Mencuccini, M., García-Valdés, R., Nadal-Sala, D., Sabaté, S., Martin-StPaul, N., Morin, X., D'Adamo, F., Batllori, E., and Améztegui, A.: MEDFATE 2.9.3: a trait-enabled model to simulate Mediterranean forest function and dynamics at regional scales, Geoscientific Model Development, 16, 3165–3201, https://doi.org/10.5194/gmd-16-3165-2023, 2023.
Granier, A., Biron, P., and Lemoine, D.: Water balance, transpiration and canopy conductance in two beech stands, Agricultural and Forest Meteorology, 100, 291–308, https://doi.org/10.1016/S0168-1923(99)00151-3, 2000.
Hoek van Dijke, A. J., Herold, M., Mallick, K., Benedict, I., Machwitz, M., Schlerf, M., Pranindita, A., Theeuwen, J. J. E., Bastin, J.-F., and Teuling, A. J.: Shifts in regional water availability due to global tree restoration, Nat. Geosci., 15, 363–368, https://doi.org/10.1038/s41561-022-00935-0, 2022.
Llorens, P., Latron, J., Álvarez-Cobelas, M., Martínez-Vilalta, J., and Moreno, G.: Hydrology and Biogeochemistry of Mediterranean Forests, in: Forest Hydrology and Biogeochemistry: Synthesis of Past Research and Future Directions, edited by: Levia, D. F., Carlyle-Moses, D., and Tanaka, T., Springer Netherlands, Dordrecht, 301–319, https://doi.org/10.1007/978-94-007-1363-5_14, 2011.
Mastrotheodoros, T., Pappas, C., Molnar, P., Burlando, P., Manoli, G., Parajka, J., Rigon, R., Szeles, B., Bottazzi, M., Hadjidoukas, P., and Fatichi, S.: More green and less blue water in the Alps during warmer summers, Nat. Clim. Chang., 10, 155–161, https://doi.org/10.1038/s41558-019-0676-5, 2020.
Qubaja, R., Amer, M., Tatarinov, F., Rotenberg, E., Preisler, Y., Sprintsin, M., and Yakir, D.: Partitioning evapotranspiration and its long-term evolution in a dry pine forest using measurement-based estimates of soil evaporation, Agricultural and Forest Meteorology, 281, 107831, https://doi.org/10.1016/j.agrformet.2019.107831, 2020.
Simonin, K., Kolb, T. E., Montes-Helu, M., and Koch, G. W.: The influence of thinning on components of stand water balance in a ponderosa pine forest stand during and after extreme drought, Agricultural and Forest Meteorology, 143, 266–276, https://doi.org/10.1016/j.agrformet.2007.01.003, 2007.

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.

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:

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

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.

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 S7-12). 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.

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)."

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 S7-12, S15 previously referenced).

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.