Response 1 to Reviewer of Manuscript: Afforestation impacts on terrestrial hydrology insignificant compared to climate change in Great Britain

Comments/Text of the Reviewer is in **black**, our response is in **blue**.

It is important to understand the relative effects of afforestation and climate on the hydrology of river catchments. This manuscript considers the effects of both afforestation and climate change in detail for catchments in Great Britain and concludes that afforestation impacts are insignificant compared to climate change. The modelling work seems to have been carried out well and the manuscript is generally well written. However what I feel is missing is a comparison with measured data on the effect of afforestation, which is needed in order to check the results are realistic and that the conclusion is valid. The authors even state on L71 "LSMs should therefore quantify projected hydrological changes whilst modellers determine if outputs are realistic". My experience suggest the conclusions are valid but it is important to demonstrate that this true.

We thank the Reviewer for their kind comments on the modelling work and writing. We respond to the second part of this paragraph in response to the point below.

JULES LSM model simulations were carried out for 51 river catchments in the UK a range of different afforestation and climate scenarios. My understanding is that the baseline model was calibrated and validated against streamflow and soil moisture data. Then afforestation and climate simulations are carried out. The results show that "Afforestation across Great Britain moderately reduces average river flow by 0.17 % PPPoA over the year" where PPPoA is the per percentage point of afforestation. Is this value realistic? In my view all the results and conclusions rely on the model producing a realistic value here and there does not seem to have been any attempt to validate this value. I appreciate obtaining good data for this is difficult, but there are UK studies available. In the wetter parts of the country there are sites at Coalburn, Plynlimon, Balquhidder and in the drier parts at Blackwood, Alice Holt, Thetford Forest and Clipstone Forest. A lot of these consider a change from grassland to coniferous forests rather than broadleaf forests considered here. Also in many cases there are point scale measurements rather than changes in catchment river flows, but they will give an idea if the results are roughly correct. For example in the Coalburn catchment (Birkinshaw et al. 2014), which has already been cited, a change from grassland to a mature conifer forest (90% of the catchment) has produced a reduction in streamflow of around 350mm for an average river flow of around 900mm, which by my calculation reduces average river flow by 0.35 % PPPoA.

It could also be argued that there is no validation that the change in climate is also producing realistic results in the model. But I am not sure how this can be checked expect that the current validation against existing streamflow measurements covers "known floods and drought events".

In a revised manuscript we would substantially revise our description of the validation of the JULES model configuration used here, drawing on an extensive range of prior studies that have evaluated the model's performance in relation to both hydrology, soil hydraulics, land cover, and biogeochemistry. In this study we require a physically, and not catchment, based model to explore the hydrological response to changes in land cover and climate across the country. JULES' strength comes from the combination of processes and systems, and not just singular ones, thus enabling exploration into the interaction between systems (such as dynamic vegetation and hydrology). This does however make it practically impossible to validate every output, particularly as there is no such data at this spatial and temporal resolution across the country (a point to be explored further below).

Importantly, this work builds upon validation and the studies of others to undertake one of the most advanced studies to date in this field. The biomechanics of JULES has been continuously improved and utilised in this study. To note the work of Harper et al. (2016, 2021) that both validated and improved the plant functional types used in JULES and found that in temperate regions that JULES produced good agreement with latent energy fluxes. The initial implementation of plant hydrology mechanics within JULES can be found in Alton et al. (2009) and shows that global runoff agrees fairly well with this implementation as well as evapotranspiration. Van den Hoof et al. (2013) and Blyth et al. (2019) also illustrate that JULES gets the broad pattern of evapotranspiration changes correct. Martinez-de La Torre et al. (2019a) demonstrate further that land surface models with vegetation dynamics generate better drydown characteristics than large scale hydrological models.

Considering streamflow and runoff, the model configuration used is that developed by Martinez-de La Torre et al. (2019b) that achieve an NSE score of over 0.8 for the River Thames. Older versions of JULES have also been shown to get the broad structure of hydrological events and extremes correct (e.g. Prudhomme et al. 2011; Harding et al. 2014). More recent version implementations of JULES have produced even more accurate results (e.g. Lewis & Dadson 2021; Mathison et al. 2023).

The main model configuration used in this study can be found in Buechel et al. (2022) as stated on line 152. Buechel et al. (2022) provides a validation of streamflow, soil moisture and evaporation (including at Alice Holt) and these can be found in the methodology section of that study. At the broadleaf woodland sites there was both systematic over and underprediction of soil moisture with an overestimation of evaporation. However, the system responses were broadly correct.

The Reviewer hits the main scientific quandary: there is no way to validate something that has yet to occur (widespread afforestation across the UK). The studies cited by the Reviewer are all coniferous afforestation and at small scales (often less than < 5 km^2). Here, in contrast, the scenarios utilised are all broadleaf afforestation at countrywide/ large catchment scales. Processes that are relevant at small scales, such as forest management strategy and ditching, will not be explicitly represented within this form of model. Neither would it be expected that the model should include these processes, particularly as the observational studies do not necessarily agree on the direction of change (e.g. Stratford et al. 2017). JULES should not be treated as completely free of uncertainty , but it broadly produces the right result for the right reasons and most importantly, has a high degree of fidelity. Precisely calibrating one output (such as streamflow) would compromise other processes within the model domain. This study provides a potential response of countrywide hydrology to afforestation.

We suggest adding a paragraph in the methodology section, providing details on validation studies that support the use of JULES to study land cover and hydrology changes.

Alton, P., Fisher, R., Los, S., and Williams, M. (2009),Simulations of global evapotranspiration using semiempirical and mechanistic schemes of plant hydrology, _Global Biogeochem. Cycles, 23, GB4023, doi:[10.1029/2009GB003540](https://doi.org/10.1029/2009GB003540

Blyth, E.M., Martinez-de la Torre, A. and Robinson, E.L., 2019. Trends in evapotranspiration and its drivers in Great Britain: 1961 to 2015. Progress in Physical Geography: Earth and Environment, 43(5), pp.666-693.

Buechel, M., Slater, L. and Dadson, S., 2022. Hydrological impact of widespread afforestation in Great Britain using a large ensemble of modelled scenarios. Communications Earth & Environment, 3(1), p.6.

Harding, R.J., Weedon, G.P., van Lanen, H.A. and Clark, D.B., 2014. The future for global water assessment. Journal of Hydrology, 518, pp.186-193.

Harper, A. B., Cox, P. M., Friedlingstein, P., Wiltshire, A. J., Jones, C. D., Sitch, S., Mercado, L. M., Groenendijk, M., Robertson, E., Kattge, J., Bönisch, G., Atkin, O. K., Bahn, M., Cornelissen, J., Niinemets, Ü., Onipchenko, V., Peñuelas, J., Poorter, L., Reich, P. B., Soudzilovskaia, N. A., and Bodegom, P. V.: Improved representation of plant functional types and physiology in the Joint UK Land Environment Simulator (JULES v4.2) using plant trait information, Geosci. Model Dev., 9, 2415–2440, https://doi.org/10.5194/gmd-9-2415-2016, 2016.

Harper, A. B., Williams, K. E., McGuire, P. C., Duran Rojas, M. C., Hemming, D., Verhoef, A., Huntingford, C., Rowland, L., Marthews, T., Breder Eller, C., Mathison, C., Nobrega, R. L. B., Gedney, N., Vidale, P. L., Otu-Larbi, F., Pandey, D., Garrigues, S., Wright, A., Slevin, D., De Kauwe, M. G., Blyth, E., Ardö, J., Black, A., Bonal, D., Buchmann, N., Burban, B., Fuchs, K., de Grandcourt, A., Mammarella, I., Merbold, L., Montagnani, L., Nouvellon, Y., Restrepo-Coupe, N., and Wohlfahrt, G.: Improvement of modeling plant responses to low soil moisture in JULESvn4.9 and evaluation against flux tower measurements, Geosci. Model Dev., 14, 3269–3294, https://doi.org/10.5194/gmd-14-3269-2021, 2021.

Lewis, H.W. and Dadson, S.J., 2021. A regional coupled approach to water cycle prediction during winter 2013/14 in the United Kingdom. *Hydrological Processes*, *35*(12), p.e14438.

Martínez-de la Torre, A., Blyth, E. M., and Weedon, G. P.: Using observed river flow data to improve the hydrological functioning of the JULES land surface model (vn4.3) used for regional coupled modelling in Great Britain (UKC2), Geosci. Model Dev., 12, 765–784, https://doi.org/10.5194/gmd-12-765-2019, 2019a.

Martínez-de la Torre, A., Blyth, E.M. and Robinson, E.L., 2019b. Evaluation of drydown processes in global land surface and hydrological models using flux tower evapotranspiration. Water, 11(2), p.356.

Mathison, C., Burke, E., Hartley, A.J., Kelley, D.I., Burton, C., Robertson, E., Gedney, N., Williams, K., Wiltshire, A., Ellis, R.J. and Sellar, A.A., 2023. Description and evaluation of the JULES-ES set-up for ISIMIP2b. Geoscientific Model Development, 16(14), pp.4249-4264.

Prudhomme, C., Parry, S., Hannaford, J., Clark, D.B., Hagemann, S. and Voss, F., 2011. How well do large-scale models reproduce regional hydrological extremes in Europe?. Journal of Hydrometeorology, 12(6), pp.1181-1204.

Stratford, C., Miller, J., House, A., Old, G., Acreman, M., Duenas-Lopez, M.A., Nisbet, T., Burgess-Gamble, L., Chappell, N., Clarke, S. and Leeson, L., 2017. Do trees in UK-relevant river catchments influence fluvial flood peaks?: a systematic review. https://core.ac.uk/download/pdf/96704761.pdf

Van den Hoof, C., Vidale, P.L., Verhoef, A. and Vincke, C., 2013. Improved evaporative flux partitioning and carbon flux in the land surface model JULES: Impact on the simulation of land surface processes in temperate Europe. Agricultural and forest meteorology, 181, pp.108-124.

Specific Comments

L85. I note that in sections 3 and 4 that the sub-sections correspond to the 3 research questions. I might be worth highlighting this here as it does make the manuscript easier to read and I did not realise to start with.

We thank the Reviewer for this helpful comment, and we can amend the manuscript accordingly to change the subtitle headings to emphasise that they relate to the 3 research questions.

L161. “JULES runs at a numerical timestep of 30 minutes”. This is clear, but what is the timestep of the meteorological input data and the streamflow data (I might have missed this)?

We are grateful to the Reviewer for highlighting the omission of this information. CHESS-met is given at a daily resolution and is disaggregated by JULES. The streamflow data calculated and compared to is at a daily resolution as well. This can be clarified within the manuscript.

L188. There should be a comparison with other hydrological models of UK catchments here, for example Lane et al. (2019) and Lees et al. (2021). These use NSE as an objective function rather than KGE, but I note in the supplementary material that NSE values are calculated. Also are you using hourly or daily discharge for the comparison?

We thank the Reviewer for raising this point and including a reference to our earlier work. Although it is beyond the scope to run more hydrological models to compare the model to, we will cite these references for readers to consult. Those studies also use KGE. It is heavily debated which calibration factor is the most appropriate (e.g. Knoben et al. 2019). The daily discharge is used and can be clarified.

Knoben, W.J., Freer, J.E. and Woods, R.A., 2019. Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. Hydrology and Earth System Sciences, 23(10), pp.4323-4331.

L199. Which Dee catchment? There are at least two rivers called Dee in the UK

We apologise for the confusion. In this context we are talking about the Dee UKCP18 river basin region, and this can be clarified within the text.

L217 “twenty UKCP18 river basin boundaries”. Why are there 20 here but in L196 there were 51?

This is because this line refers to the ‘UKCP18 river basin boundaries’ whereas the 51 are for the catchments as written. We can emphasise this difference within the manuscript if required.

L233. “Proportional Influence of Afforestation Compared to Climate”. You talk about changing precipitation, temperature, and CO2 (i.e three variables) but there are also changes in afforestation (so a fourth variable). I got confused and had to go back and read the section again.

We apologise for the confusion and will rewrite the section for clarity.

L288-292. This could do with being rewritten. The authors define soil evaporation “Simulated soil evaporation, including both evaporation from the soil surface and plant transpiration” but then consider transpiration separately but not soil surface evaporation separately.

These lines can be rewritten as there is no way to disentangle soil evaporation and soil transpiration. This line was written incorrectly and was meant to be modelled stomatal conductance and not transpiration.

L308 Table 1. How are “Soil moisture” and “Canopy Storage calculated”. Is this the average over the simulation period? The data in the “future” columns is discussed in section 3.3, can the caption be changed to make this clear, when I first read this I got confused about why these results had suddenly appeared.

They are calculated in the same way the other variables are calculated. The table subtitle is now written ‘Total Column Soil Moisture’ to clarify that the soil moisture is for the whole column. The caption can be altered to explain for the references.

L312 Table 2. The correlations seems to be calculated for the percentage change. I would be interested in calculating the correlation for the absolute change. As I am sure the authors are aware many catchments in Scotland and North West England and Wales have precipitation totals that are around 3 times those in South East England. So if they have a similar percentage change they will have an absolute change that will be very different, which suggests there will be a significant influence due to the location for absolute change.

Flows across the catchments are different, as the Reviewer notes, however we need the changes to be on a comparative scale to observe the relative effect of woodland on flow percentiles. The use of a percentage scale enables us to deduce the relative impact of afforestation on hydrological processes across catchments of different sizes and hydrological magnitudes. Fundamentally the processes occurring in one area are the same elsewhere in the model domain and so use of the percentage change can be used to calculate the absolute change if needed (but could not be achieved the other way around). As shown in Table 2 and as written in the manuscript, location has a significant impact on the response of high river flow in comparison to woodland. If absolute values were used, the scale of the catchment and location would have a far greater impact on flow generation and not enable the influence of woodland to be ascertained (which is the question being asked here).

L325 Figure 4. In winter canopy evaporation increases as expected (Page et al. 2020). But what is driving the massive increase in soil evaporation. Soil evaporation “includes both evaporation from the soil surface and plant transpiration” but L291 says transpiration decreases with afforestation. This implies that increasing the forest is producing a massive increase in evaporation from the soil surface in winter. This does not make sense to me as a mature forest will have very little soil surface evaporation in winter. So either the model is doing something strange or I have misunderstood the results.

In winter there is diminished canopy protection due to the phenology implementation in the model, which leads to an increase in soil evaporation. On line 291, transpiration was the wrong word and was meant to be ‘stomatal conductance’ and will be changed in the manuscript. Further information on the evaporation effect can be found in Buechel et al. 2022 but it is due to the change from grasslands to woodland. To summarise a few points from that manuscript:

- Soil evaporation increased due to broadleaf woodland losing leaves. Reduced canopy cover increases soil exposure to incoming short-wave radiation and reduces aerodynamic resistance, which increases potential evaporation.
- In summer months a decrease in modelled soil evaporation increases topsoil moisture (because the increase in canopy foliage reduces soil exposure to short-wave radiation and surface wind speed).

- Topsoil moisture may increase because of root structure differences in broadleaf woodland and grasslands representation in JULES.

Lines 457-458 also explain that evaporation is likely to be overestimated in the model.

Buechel, M., Slater, L. and Dadson, S., 2022. Hydrological impact of widespread afforestation in Great Britain using a large ensemble of modelled scenarios. *Communications Earth & Environment*, *3*(1), p.6.

L340 Figure 5 In the bottom panel the runoff increases slightly and the canopy evaporation decreases slightly for 100% afforestation compared to 50% afforestation. The changes are small but I was wondering why this was happening.

This is likely due to means of the variables, runoff and canopy evaporation, being calculated and so the extreme increase in subsurface runoff being calculated for 100% afforestation is leading to that increase. For canopy evaporation, again there is such a spread that it is statistically insignificant when compared to other factors.

L360 Figure 6. I feel it would be easier to interpret if the y axis had the same scales for Top Flows, Median Flows and Low Flows

We understand the Reviewer's perspective on this and we initially presented the data with the same scale. However, it makes the data more difficult to interpret (particularly with the larger y axis range for the Median Flows). In Figure 6, we have used the same horizontal lines for comparison.

L406-L407 "Average simulated river flow, compared to present, drops at a slightly lower rate of -0.12 % PPPoA (ρ = -0.82) [Table1]. Runoff decreases with afforestation at a comparable rate to present (-0.27 %; 1.84 mm yr-1 PPPoA)". I do not remember seeing the difference between "river flow" and "runoff" defined. Can this be added somewhere and then explained why they are getting different results. Also Table 1 has runoff but no river flow. where is the value of -0.12% in Table 1?

Model definitions of river flow and runoff can be included. River flow is that taken at the gauging station location after the runoff has been routed through the kinematic wave River Flow Model, whereas runoff is taken as the specific combination of subsurface and surface runoff for the entire catchment.

River flow was removed from the table in editing and this figure reference can be removed. We thank the Reviewer for spotting this error.

L421 Discussion. The discussion is interesting but my personal feeling is that it is on the long side (I struggled to concentrate whilst reading it) and there are bits that are not completely relevant.

We are sorry that the Reviewer believes that there are irrelevant parts to the discussion. However, we have included everything we believe is necessary and this work is trying to remain of broad interest to those working on afforestation policy, woodland hydrology, and modellers in both hydrology and land surface models. It is therefore difficult to write a discussion that is relevant to just one group and thus the discussion is aimed at all those who might have an interest on this topic. However, we appreciate the Reviewer's point and we will attempt to streamline the text so that it does not feel unnecessarily long.

L432 "suggesting potential reductions in water yield can be directly estimated from the areal extent of woodland planted rather than its location". This follows from an earlier point this may be true for the percentage reductions but does the location affect the absolute reductions?

The trends in the absolute reductions are the same as the percentage reductions. Usage of percentage enables the changes to be on a relative scale between catchments and regions. Therefore, one can broadly convert the percentage changes into absolute changes.

L466 "In reality, tree root depths would be much deeper" is this due to using on a 3m deep soil column in JULES? Maybe make this clear.

We will edit the text to clarify, e.g. 'In reality, tree root depths would be much deeper than currently represented in JULES…'.

L485 "Therefore in reality, afforestation may have a muted influence on streamflow in these regions with roots accessing the deeper groundwater (Roberts and Rosier, 2005)." I do not understand this. If forests can access deeper groundwater then it might have a greater influence on streamflow in the longer term as it can transpire water even when there is a meteorological drought.

We understand the confusion of this and can replace 'muted' with 'more subtle and greater' to emphasise the fact of woodland impacting groundwater.

L513 "afforestation both decreases and increases streamflow". This needs explaining

We can add '… depending on catchment and antecedent conditions' to provide greater clarity.

L520-L527 Could this bit be removed? It is all a bit vague I do not really understand the point it is making.

We understand that the Reviewer may not be interested in this, but we believe it should remain to benefit the hydrological and land surface modelling communities. However, it can be edited for greater clarity if there is confusion.

L560 where is Figure 8?

We thank the Reviewer for finding this error from editing. The actual figure caption for this should have been Figure 6.

L580 "with climate…". Is this bit a repeat from the previous section?

This is not a repeat of the previous section. This is to highlight to the reader that if considering the role afforestation may have on flood processes, there is unlikely to be a significant impact of woodland being able to reduce these high flows.

L585 "By applying atmospheric changes across the whole country, variations in landcover, topography and soil type are insufficient to substantially alter the hydrological response." Has this been shown for topography and soil type?

Although there has been no explicit testing for soil and topography type, the alterations in hydrological processes are more substantial with atmospheric changes (e.g. precipitation) than due to the differences in land surface properties of the different regions. We have tested across a wide range of topographies and soil types with the number of catchments and find that the effect of land-cover on the response is much lower than the effect of changes in climate. Therefore, it can be concluded that within this model setup the driving atmospheric data is more important for generating hydrological responses. The lines 586-589 emphasise that this could be due to overparameterization in JULES or missing relevant processes.

References

Birkinshaw, S. J., Bathurst, J. C., & Robinson, M. (2014). 45 years of non-stationary hydrology over a forest plantation growth cycle, Coalburn catchment, Northern England. *Journal of Hydrology*, *519*, 559-573.

Lane, R. A., Coxon, G., Freer, J. E., Wagener, T., Johnes, P. J., Bloomfield, J. P., ... & Reaney, S. M. (2019). Benchmarking the predictive capability of hydrological models for river flow and flood peak predictions across over 1000 catchments in Great Britain. *Hydrology and Earth System Sciences*, *23*(10), 4011-4032.

Lees, T., Buechel, M., Anderson, B., Slater, L., Reece, S., Coxon, G., & Dadson, S. J. (2021). Benchmarking data-driven rainfall–runoff models in Great Britain: a comparison of long short-term memory (LSTM)-based models with four lumped conceptual models. *Hydrology and Earth System Sciences*, *25*(10), 5517-5534.

Page, T., Chappell, N. A., Beven, K. J., Hankin, B., & Kretzschmar, A. (2020). Assessing the significance of wet-canopy evaporation from forests during extreme rainfall events for flood mitigation in mountainous regions of the United Kingdom. *Hydrological Processes*, *34*(24), 4740-4754.

We would like to thank the Reviewer again for the time and effort they put into this review and we look forward to implementing their suggestions to improve the manuscript.