General comments

The manuscript presents a coupled FORMIND–mHM framework that integrates dynamic forest processes with catchment-scale hydrological modelling. This development has the potential to improve the representation of vegetation–hydrology interactions, particularly in forest-dominated catchments. However, several aspects of the coupling implementation and model evaluation require further clarification and analysis.

My main concern is that the manuscript places considerable emphasis on describing the coupling framework, while the evaluation does not yet sufficiently demonstrate its added value relative to the standalone models. In particular, the scale-transfer mechanisms for hydrological processes between FORMIND patches and mHM grid cells require further clarification.

Furthermore, the model evaluation relies on limited observations and lacks sufficient process-based analysis to explain how and why coupling affects gross primary productivity (GPP), evapotranspiration (ET), discharge, and water-balance partitioning.

A clearer explanation of the coupling mechanisms, together with a more comprehensive assessment of their effects on vegetation and hydrological processes, would substantially strengthen the manuscript.

> Thank you very much for your detailed review of our manuscript and all the helpful suggestions and ideas! Below, we respond to each of your comments in detail and outline how we intend to address them in the revised manuscript.

Detailed comments

2.1 Forest model FORMIND

Line 125: The authors should briefly explain how LAI is calculated in FORMIND or provide an appropriate reference.

> The LAI of an individual tree crown is calculated as an allometric function of its stem diameter. The LAI of a stand is derived from the leaf areas of all the trees within the stand, distributed across the vertical crown layers. Low values of GPP can lead to a decrease in LAI, too.
>
> We will add Fischer at al. (2016) as a reference for the LAI calculation in FORMIND in Sect 2.1.

2.2 Mesoscale hydrologic model mHM

Line 140: The manuscript indicates that mHM uses prescribed LAI. The authors should clarify whether this LAI is retained or replaced by FORMIND-simulated LAI in forested areas after coupling.

> In forested areas, the LAI is indirectly replaced, as LAI-based processes are simulated by FORMIND in the coupled framework. In all non-forested areas, the LAI is still needed for the calculation of effective precipitation.
>
> We will clarify this in the revision in Sect 2.3.

The authors should also clarify whether LAI is still required for the hydrological calculations in non-forested areas and, if so, how it is prescribed or calculated. Potential inconsistencies arising from the different LAI representations in forested and non-forested areas should also be discussed.

> In the coupled framework, FORMIND models the vegetation only in forested areas. In the non-forested fraction of each cell, mHM requires an LAI input for calculating

evaporation as in stand-alone mHM. In the non-forested areas, we use the same prescribed monthly LAI in both configurations. The non-forested areas are therefore treated identically in FORMIND-mHM and stand-alone mHM, and differences between both configurations arise only through the forested fraction. We verified that the non-forested output of FORMIND-mHM matches stand-alone mHM in cells without forest.

We will clarify this in Sect. 2.3.

Lines 142–143: The authors should briefly describe the main processes or governing equations used to simulate soil-water dynamics in mHM, or provide appropriate references.

Thank you, this is a good point. Soil moisture dynamics are expressed as $x_i^l = (1 - p^l)\, I_i^{l-1}(t) - E_{3i}^l(t) - I_i^l(t)$, with p^l being the overall influx fraction accounting for impervious cover within a cell, *I* being Recharge (effective precipitation or infiltration intensity, depending on soil layer i), and E_3 being actual evapotranspiration.

We will add this in the revision in Sect. 2.2.

2.3 Coupling components and implementation

The implementation of LAI and soil-water processes across the coupled components remains unclear. The authors should explain how FORMIND-simulated forest dynamics influence hydrological calculations, whether these effects are mediated solely through the forest exfiltration flux (exfilforest), and whether the original LAI-related processes in mHM remain active. The authors should also clarify how differences in soil-layer configurations between the two models are handled.

Thank you. In forested areas, FORMIND-simulated forest dynamics influence soil moisture dynamics directly through the plants' water demand, which is calculated in FORMIND. The LAI influences plants' water demand as the demand is proportional to the GPP, which in turn is LAI dependent. The soil module checks whether such demand can be fulfilled with current soil moisture levels. Here, original LAI-related processes in mHM are inactive.

Non-forested areas are independent from the coupling mechanism. Therefore, original LAI-related processes in mHM remain active.
There is no vertical exchange between the cells at the upper-soil level. Meaning, forested and non-forested areas are only aggregated at the catchment-scale as their respective exfiltration in the lower soil are merged by areal fraction. We will elaborate on this.

Differences in soil-layer configurations between the models have not been considered, as the coupled framework only uses one soil moisture model. Here, we follow the German Drought Monitor.

We will clarify this in Sect. 2.3.

The origin and implementation of the upper-soil component require further explanation. Is this component independent of FORMIND and mHM, or are the soil-water and snow-process formulations from mHM implemented within FORMIND? If the latter, the authors should explain how these processes are transferred between the different spatial scales. In particular, the procedures for transferring information from mHM grid cells to FORMIND patches and aggregating patch-level outputs back to the grid scale should be described. The representation of forest disturbances, particularly wildfire, should also be clarified.

Thank you for this comment.

The upper-soil component is implemented as an independent python component acting on a 20 m x 20 m grid size. The functional behaviour of this component is identical to mHM, only transferred from Fortran into Python. Accordingly, the snow-

processes are implemented as a stand-alone Python compartment, too. This is necessary, as processes in mHM are calculated in a serial way, with soil moisture calculations using snow melt as input. Therefore, there are no exchanges from the 4 km x 4 km mHM cells to the 20 m x 20 m FORMIND patches, as the mHM logic has been transferred to the FORMIND scale. The forest exfiltration is then averaged from patch-level to grid-scale.

We will clarify this in the revision in Sect 2.3.

In Theory, the coupled model simulates non-drought related forest disturbances in exactly the same way as the standalone FORMIND model. The forest fire module is explained in greater detail in Fischer (2021), but is not used in this study.

We will include this as a potential future application.

Figure 1: The current schematic is too simplified to adequately illustrate the coupling architecture. The authors should identify the main components, scale-transferred process components, scaling component, and the role of FINAM, together with the variables exchanged between FORMIND and mHM and their directions of transfer.

Thank you. We will make the 'Snow Module' and the 'Upper Soil' component identifiable as the scale-transferred process components. Further, we will distinguish between forest, non-forest and catchment. We will further clarify the role of FINAM for the coupled framework.

Line 172: The repeated citations to Müller et al. (2025) when introducing FINAM could be reduced, retaining the reference at the most relevant point in the methodological description.

We will do that.

Line 185; Figure 1: The upper-soil component reportedly uses hydrological formulations from mHM to calculate ET, but Figure 1 places this component within the forest-related framework. The authors should clarify whether it calculates ET only for forested areas or for the entire catchment. If it also represents non-forested areas, Figure 1 should be revised to reflect its spatial scope.

The separate upper soil component calculates ET only for forested areas, but using the same logic and calculations as mHM which continues to calculate ET for non-forested areas. The spatial heterogeneity in soil moisture is an important driver for forest dynamics in FORMIND. Introducing the stand-alone soil component allows us to preserve the spatial heterogeneity on a 20 m x 20 m scale, while using mHM hydrological formulations.

We will rename the Soil Module to increase clarity of Figure 1.

Spatial resolution: The authors should justify the choice of a 4 km × 4 km hydrological grid and discuss the potential implications of this resolution for representing spatial heterogeneity within the catchment, particularly compared with a finer resolution such as 1 km × 1 km.

Thank you for this comment. We chose a 4 km × 4 km resolution following the setup of the German Drought Monitor, for which mHM has been extensively evaluated (Boeing et al., 2022). Several aspects limit the effect of this choice compared to a 1 km × 1 km cell size on our results. First, mHM derives its effective parameters from input data at 100 m × 100 m resolution via multiscale parameter regionalisation (MPR; Samaniego et al., 2010), which accounts for sub-grid heterogeneity in soil properties and land cover and yields parameters that are transferable across scales (Kumar et al., 2013). Further, mHM is effectively scalable, simulated water fluxes at the catchment scale do not depend on spatial resolution. And lastly, the coupled framework scales with the number of grid cells, and a 1 km grid would require about 16 times more FORMIND instances.

In the revised manuscript, we will add a sentence in Sect. 2.3 explaining the choice of grid resolution.

2.4 Model evaluation and analysis

Figure 2: In addition to forests, agricultural areas occupy a substantial proportion of the study catchment. The authors should clarify how cropland and other non-forested areas are represented in the coupled framework, particularly how their LAI is prescribed or calculated and whether crop growth and agricultural management processes are explicitly considered. The implications for catchment-scale water and carbon fluxes should also be discussed.

Thank you. All other non-forested areas are not affected by coupling and are represented by a separate mHM component that is integrated into the model framework. Here, we use a monthly LAI value, which varies over the course of a year and can be regarded as an indicator of plant growth. Management measures are not considered. There is ongoing work to include a grassland model into the framework for such areas.

In Section 2.3, as well as throughout the paper, we will make it clearer that we are using a forest gap model and are therefore not considering vegetation dynamics as a whole, but only forests.

Furthermore, we will make this clear distinction evident right from the title of the revised manuscript.

2.4.1 Model setup for the Selke catchment

Lines 227–228: The authors should provide appropriate references for the meteorological forcing datasets. A summary table of the input datasets used by each model component and the observational datasets used for evaluation would improve reproducibility.

Thank you for this suggestion. We will introduce such a table in the revised manuscript.

The spin-up procedure is described for FORMIND but remains unclear for mHM and the hydrological components. The authors should clarify how these components were initialized and whether an additional spin-up or adjustment period was performed after coupling to ensure consistent initial hydrological and vegetation conditions.

Thank you. The standalone mHM version as well as the integrated mHM versions have a spin-up period. The coupled soil module is initialized from a mHM restart file. This restart file is created after the spin-up and before the actual simulation period starts, to ensure identical storages.

We will add information on this in Sect. 2.4.

2.4.2 Stand-scale evaluation

Figure 2: The location of the Hohes Holz eddy-covariance site should be indicated in Figure 2 to clarify its spatial relationship to the catchment.

We will add the location to the Germany map in Fig. 2.

Line 241: The authors should clarify the sources and spatial representativeness of the GPP and ET observations, particularly whether the eddy-covariance measurements represent forested areas only and how their spatial footprint corresponds to the simulated variables.

We will add a sentence clarifying this in Sect. 2.4.2.

2.4.4 Model analysis

Line 267: The parameter optimization procedure requires further clarification. The authors should specify which parameters were optimized, particularly whether optimization was limited to the mHM canopy interception parameter and the FORMIND water-use efficiency (WUE) parameter. Their initial values, optimization ranges, and optimized values should be provided, preferably in a summary table with appropriate references.

> Thank you for this suggestion. We will clarify the parameter optimization procedure and add a table to the Supplementary Information including the optimized parameters, their parameter ranges as well as the initial values.

3 Results

3.1 Evaluation of the vegetation component

The evaluation of the vegetation component remains limited. The authors should further assess how coupling affects GPP and ET, including their seasonal variability, and explain the underlying processes. Furthermore, the evaluation relies on a single forest eddy-covariance site. The authors should clarify how GPP and ET in non-forested areas, particularly cropland, are represented and evaluated. If suitable observations are unavailable, this limitation should be acknowledged.

> Thank you for this point, which was also raised by the other reviewers. We will add a catchment-scale evaluation of ET and GPP against remotely sensed MODIS products for the Selke catchment. The products will be aggregated to the 4 km model grid and compared with FORMIND-mHM (and, for ET, with stand-alone mHM), at monthly resolution, for the mean seasonal cycle and annual sums. These products are themselves model-based and associated with significant uncertainties (Wang et al., 2017; Hu et al., 2015). Therefore, they constitute an important independent reference, but may not be authoritative.
>
> For non-forested areas, there is no GPP calculation, as GPP is not computed in mHM. ET is calculated according to Eq. 4. This has not been explicitly evaluated in this study, as mHM runs unchanged in non-forested areas.

3.2 Evaluation of the hydrologic component

Figure 4: The standalone mHM simulation achieves a KGE of 0.91, whereas the coupled FORMIND–mHM simulation shows a lower KGE of 0.85. The authors should explain the reasons for this decline, particularly which hydrological processes are affected by coupling.

> Thank you. The results in Fig. 4 show an mHM simulation in which the parameters have been optimised to match the measured values at the gauge Meisdorf. These include parameters that control both evaporation from the canopy and transpiration. Whilst FORMIND-mHM uses the same soil parameters, FORMIND's vegetation-specific parameters were not optimised for the Selke catchment. The figure shows that, although FORMIND-mHM can reproduce the dynamics, it does not quite match the intensity of the extreme events. As evaporation and transpiration are the processes affected, the partitioning of water fluxes shifts from an evaporation-dominated to a transpiration-dominated pattern, leading to a reduction in soil moisture and, consequently, a decrease in runoff volumes. Nevertheless, both values represent a very good KGE.
>
> We will clarify this in Sect. 4.2 and further add that the goal is not to outperform either model, but to show that the coupled framework is usable in both domains, generating respectable results.

The evaluation should extend beyond time-series comparisons to examine seasonal differences between the standalone and coupled simulations, particularly in spring and

summer. Additional analyses, such as seasonal cumulative distribution functions (CDFs) of soil saturation and ET, could help characterize changes in hydrological conditions following coupling.

> Thank you; this is a very good suggestion. We will include CDFs in the revised manuscript.

Furthermore, given that the catchment includes three discharge stations, the authors should consider evaluating discharge at the catchment outlet (Hausneindorf) to assess model performance beyond the forest-influenced Meisdorf subcatchment.

> We deliberately decided not to use the gauge Hausneindorf for assessing model performance. For non-forested areas, mHM continues to run as before. The model has been comprehensively evaluated and assessed, and we have verified that the results from FORMIND-mHM for non-forested areas in cells without forest correspond to those of the standalone mHM. As there is barely forest in the lower Selke catchment, added value for assessing the coupled framework is limited. Naturally, differences between FORMIND-mHM and standalone mHM are proportional to the forested area.
>
> We will clarify this in Sect. 2.4.3.

3.3 Water balance partitioning and effect of parameter optimization

The water-balance analysis should provide a more detailed interpretation of how coupling affects the individual hydrological components, rather than primarily reporting changes in performance statistics.

> Thank you for this suggestion. We agree that the water-balance analysis focused too strongly on performance statistics. In the revised manuscript, we will extend Sect. 4.1 to discuss how the coupling affects the individual water-balance components, namely evapotranspiration (separated into transpiration, interception and evaporation from canopy), soil water storage, and exfiltration. In particular, we will explain the mechanisms behind the observed changes, how dynamic forest vegetation simulated by FORMIND alters transpiration and thereby soil moisture and runoff generation in mHM.

The authors should quantify changes in individual water-balance components, including ET, runoff, and soil-water storage changes, and explain the processes responsible for these differences. A comparative figure showing the proportions of individual water-balance components in standalone mHM and coupled FORMIND–mHM simulations would help illustrate how water-balance partitioning differs between the two scenarios.

> Thank you, we will add a comparative figure to the Supplement Material of the revised manuscript.

3.4 Parameter sensitivity tests

Line 332: Correct a punctuation error in the WUE parameter description ("WUE, =, 8").

> Will be corrected.

Figure S4: The authors should clarify which of the four parameter optimization runs is presented in Figure S4 and explain whether only one run is shown because the results are similar across all four runs.

> Thank you. S4 presents the results for four runs with an unoptimized mHM just varying the canopy interception constant.
>
> We will clarify this in the Figure description.

The sensitivity analysis considers only two parameters, one of which appears to have limited influence on the evaluated model responses. The authors should justify the selection of these parameters and discuss whether other parameters may have a greater influence on vegetation and hydrological processes. Sensitivity across different sub-catchments could also be considered.

Furthermore, the extent to which parameter optimization contributes to the added value of the coupling framework remains unclear and should be discussed.

> Thank you for this comment. We agree that the rationale behind the parameter selection was not sufficiently explained. The analysis was not intended as a comprehensive sensitivity analysis of FORMIND or mHM, but as a case study. We selected water-use efficiency (WUE) and the interception constant because each represents a process from the domain of the respective other model: WUE governs the vegetation water demand passed from FORMIND to mHM, while interception controls the water reaching the soil and thus the soil water stress passed from mHM to FORMIND. Other parameters may influence vegetation and hydrology more strongly in general, but are not directly related to the coupling interface. The limited influence of the interception constant indicates that the stand-alone mHM responds only weakly to canopy changes. We interpret this as a structural limitation of the stand-alone model rather than as evidence that canopy dynamics are hydrologically unimportant, and it is precisely where the coupling adds value by linking plant water demand directly to the simulated canopy structure.
>
> In the revised manuscript, we will clarify this motivation in 2.4.2 and refer to the analysis as an exploratory sensitivity test.

3.5 GPP–discharge relationship

Figure S8: The three discharge stations exhibit similar GPP–discharge correlation patterns. The authors should clarify whether this similarity is related to the relatively small catchment area and potentially limited spatial variability in meteorological forcing. They should also discuss whether the discharge records at these stations exhibit similar temporal variations despite differences in discharge magnitude, which may explain the consistent GPP–discharge relationships.

Supplementary figures: Figures S1, S3, and S5–S7 do not appear to be explicitly referenced or discussed in the main text. The authors should ensure that these figures are appropriately cited and briefly explain their relevance to the analysis.

> We will add this in the revised manuscript.

4 Discussion

Line 376: The description of LPJ-GUESS lacks appropriate references. LPJ-GUESS simulates vegetation dynamics through plant functional types (PFTs) and age cohorts or individuals within replicate patches, explicitly representing competition for resources within patches. Therefore, the authors should revise their statement regarding the lack of fine-scale competition in LPJ-GUESS.

> Thank you for pointing this out. We will add Sitch et al. (2003) and Hickler et al. (2012) as references when describing the vegetation representation of LPJ-GUESS in the revised manuscript. LPJ-GUESS explicitly represents competition for resources among plant functional types and cohorts or individuals within replicate patches, and we will revise our statement accordingly. However, these patches are spatially independent statistical replicates without horizontal position, and individuals within a cohort are identical. The difference to FORMIND therefore lies not in whether competition is represented, but in how explicitly vegetation structure is resolved: FORMIND simulates every tree individually on spatially arranged patches, allowing interactions among neighboring patches.

We will clarify this distinction in the revised text in Sect. 4.

Furthermore, the comparison with other models should be substantially shortened, allowing the Discussion to focus more on the coupling mechanisms, findings, and limitations of the present study.

We will shorten the comparison in Sect. 4.

4.1 Implications of coupled vegetation–hydrology modelling

Lines 395–398: The authors should consider including a spatial evaluation of soil moisture using available gridded observational or remote-sensing datasets. Where possible, comparing both standalone mHM and coupled FORMIND–mHM simulations against these observations would help assess whether coupling improves the spatial representation of soil-water availability.

Thank you, this is a good point. For soil moisture, no suitable independent product is available at the scale of our catchment (458 km²): satellite products have a coarse resolution and represent only the uppermost soil layer, whereas the simulated saturation refers to the full soil column. The German Drought Monitor is itself based on mHM. Therefore, we just use stand-alone mHM, whose soil moisture was evaluated against in situ observations in Germany (Boeing et al., 2022), as a reference for the dynamics that the coupling is not expected to change (seasonal cycle, timing).

Lines 414–419: The interpretation of the GPP–discharge relationship should be more cautious. The authors suggest that the seasonal transition from negative to positive correlations reflects the onset of water limitation. However, this conclusion cannot be established solely from the GPP–discharge relationship in a single catchment.

Thank you. We will include a time series analysis of the soil stress factor in the revised manuscript. This could offer further insights, if the described emerged pattern could be driven by the onset of water limitation.

The seasonal variations in precipitation, ET, and vegetation phenology may also contribute to this pattern. The authors should provide additional evidence from precipitation, ET, and soil-moisture dynamics to support their interpretation and avoid presenting the findings as a reversal of previous studies without sufficient evidence.

We agree that precipitation and soil moisture may act as common drivers of both GPP and streamflow. Since ET and phenology are tightly coupled to GPP in FORMIND, we focus on precipitation and soil moisture. In the revised manuscript, we will 1) add the corresponding correlations of precipitation and soil moisture with GPP and Q, and 2) compute partial correlations between GPP and Q controlling for precipitation and soil moisture. We will also revise the discussion.

4.2 Trade-offs, applicability, and technical considerations

The authors should further discuss the limitations of the coupled framework, particularly the uncertainties introduced by spatial aggregation between FORMIND patches and mHM grid cells. The potential inconsistencies arising from different vegetation representations, particularly for agricultural and other non-forested areas, should be discussed.

Furthermore, given that the framework is evaluated in only one catchment with limited discharge observations, its applicability to other catchments, particularly those dominated by agricultural or non-forest vegetation, remains uncertain and should be discussed.

Thank you very much for pointing this out. It is true that an assessment based on a single catchment area allows only cautious conclusions to be drawn about other catchment areas and should be expanded upon in future work. In Section 4.2, we discuss the advantages and disadvantages of coupling itself, as well as the

application of the coupled framework. There, we want to point out that for catchment areas used for agriculture or not dominated by forest, the coupled framework does not contribute a lot of added value, as the strength of coupled framework does not apply in these cases.

We will clarify in Sect. 4.2, as well as throughout the paper, that we do not consider all vegetation dynamic, just forests.

5 Conclusion

The Conclusions should emphasize the integration of dynamic forest processes into distributed hydrological modelling rather than implying that the framework explicitly represents dynamic vegetation processes across all land-cover types. The authors should also acknowledge the main limitations related to spatial scaling, non-forest vegetation representation, and the extent of hydrological improvements resulting from coupling.

Thank you. We will clarify that the framework represents dynamic forest processes. Whilst the representation of forest areas has improved compared with the standalone mHM, the representation of non-forest areas remained unchanged. We would therefore argue that this does not constitute a limitation. However, improving the dynamic processes of other vegetation types represents an interesting avenue for future work.

References

The references can be found in the reviewed main text. The only new reference used in our response is the following:

Fischer, R. The Long-Term Consequences of Forest Fires on the Carbon Fluxes of a Tropical Forest in Africa. Appl. Sci. 2021, 11, 4696. https://doi.org/10.3390/app11104696

Hickler, T. et al. (2012). Projecting the future distribution of European potential natural vegetation zones with a generalized, tree species-based dynamic vegetation model. Global Ecology and Biogeography, 21(1), 50–63. https://doi.org/10.1111/j.1466-8238.2010.00613.x

Hu, G. et al. (2015). Comparison of MOD16 and LSA-SAF MSG evapotranspiration products over Europe for 2011. Remote Sensing of Environment. 156, 510-526.

Sitch, S., et al. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. Global Change Biology, 9(2), 161–185.

Wang, L. et al. (2017). Evaluation of the Latest MODIS GPP Products across Multiple Biomes Using Global Eddy Covariance Flux Data. Remote Sensing, 9(5), 418. https://doi.org/10.3390/rs9050418