Response to reviewers regarding Manuscript ID egusphere-2024-2132 entitled: "Hydro-geomorphological modelling of leaky wooden dam efficacy from reach to catchment scale with CAESAR-Lisflood 1.9j" to Geoscientific Model Development.

We would like to thank Paul Quinn and the two anonymous Reviewers for their feedback on this manuscript, as well as Andy Wickert for overviewing the submission and review process. We have responded to all Reviewer comments as detailed below.

Original text from the Topic Editor and Reviews is blue.

Responses are black, Unedited text is grey, Edited text is **bold and black**.

Reviewer 1 responses to first paper revision and author comments given below in red text.

Reviewer 1 General summary

This paper presents and tests a new extension of the CAESAR-Lisflood landscape evolution model, that enables hydro-geomorphological modelling of in-stream wooden leaky barriers. The paper is interesting and well presented. It is also novel by being the first model published that considers both geomorphic and hydrological processes and the interactions between them which is of great importance for practitioners of natural flood management (NFM). The topic is also relevant to the EGU Sphere Geoscientific Model Development journal.

We thank the reviewer for their interest in the manuscript and support in acknowledging the importance of the work.

Ok.

The paper has potential as a worthy contribution of a new model tool with some interesting insights into processes and responses that could have real world implications. However, the paper would benefit from some improvements. Firstly, a more up to date review and comparison with the topic of the hydro-geomorphological functioning of leaky barriers is needed. Much progress has been made recently especially with regards to modelling the hydraulic and hydrological effects of leaky barriers but these studies are not included and may help to improve the insights given in the discussion.

Our initial aim was to remain specific to the representation of leaky dams in terms of numerical modelling, however we agree that the addition of a focused review of the functioning of leaky dams would benefit the manuscript. We have added more reference to the literature including field observations, numerical modelling and flume studies as detailed below in the line-by-line comments.

This has now been improved but see specific clarifications required below.

Secondly, a clearer research context to help justify the paper with better stated aims are needed. For example, at the moment there is little exploration of why it is important to consider geomorphic processes and what the specific geomorphic aims of this study are.

We have added this in response to line-by-line comments as detailed below.

Now improved.

Thirdly, more consideration of the validity of predictions and applicability of the tool in the real world is required. The authors state that the model is heuristic, and it is unvalidated which means its reliability for making worthwhile predictions is unknown. However, at the same time the authors advocate that the model is useful for practitioners but given the uncertainty of predictions perhaps this is not a valid viewpoint to take.

The aim of the model is to understand the relative impact of leaky dams on both the hydrology of the system and the geomorphology. As the reviewer correctly indicated in a specific comment below (referring to L477–480), there isn't the data to calibrate or validate sediment transport in the real world. Therefore, our adaptation to the CAESAR-Lisflood model in its current format is a useful addition for exploring the potential impacts of different leaky dam designs—in terms of height, gap, roughness and location—as part of a preliminary scoping study. We agree that future work is required to refine the applicability to functioning systems, however this cannot come without first addressing the data paucity surrounding leaky dams, sediment transport, and the geomorphological change they may induce. We have addressed this point throughout the line-by-line comments below.

Again, see comments below regarding this.

Specific points

L24 State recurrence interval or annual exceedance probability of storm event to give an idea of its magnitude. Also state for what catchment size.

We have adjusted this sentence as follows:

The findings show that simulating sediment transport increased the volume of water stored in the test reach **(channel length 160 m)** by up to an order of magnitude whilst reducing discharge by up to 31% during a storm event **(6 h, 1 in 10-year event)**.

Still need detail on catchment size, this important context to present.

L32-35 Low risk yes but potential for structure washout and displacement perhaps should be acknowledged.

This now reads:

NFM is becoming increasingly popular with flood risk managers due to its multiple benefits and perceived low risk**, however due to altering the hydrological regime, there is potential for structures to become displaced and washed out (e.g., Nisbet et al., 2015). NFM is also an effective method to** engage local communities and land users in potentially reducing flood risk (Burgess-Gamble et al., 2017; Dadson et al., 2017; Newson et al., 2021).

Ok this looks better.

L36 What is meant by 'river engineering'? Seems like a vague term to use.

To improve clarity, we have removed the reference to river engineering.

Ok.

L46-48 More nuance and specific reference to sources that back up these claims is needed. These benefits are often perceived and have not been quantified comprehensively or assessed.

We have adapted this sentence as follows to include recent publications in support of this:

Reintroduction of wood to the river channel is a popular form of NFM, employed for multiple co-benefits such as habitat creation and ecological enhancements (e.g., Wohl, 2017; Ockelford et al., 2024) as well as flood peak reduction (e.g., van Leeuwen et al., 2024; Villamizar et al., 2024).

But there are still deficiencies of understanding and assessment of LDs specifically which may differ to other types of large wood addition (e.g. bank attached or medial structures used to improve habitat) or naturally occurring large wood structures. For example, the number of studies that quantify empirically the flood peak mitigation of these structures is limited. Also, the study of Villamizar is based on a model whereas the van Leeuwen study is empirically based, this distinction should be made.

L50-62 A more up to date and accurate reflection on the knowledge gaps and recent advancements on understanding is needed. For example, see the work of Follett and Hankin (2022) and Geertsema et al., (2020) on approaches to model the hydraulic effects of in channel large wood interventions. Also, the work of Lo et al. (2022) gives field based observations on the geomorphic effects of leaky barriers and the work of Norbury et al. (2021) and Van Leeuwan et al. (2024) measure the hydrological effects of structures using field data. Flume based studies on hydro-geomorphic responses are also potentially useful to synthesise and compare with (e.g. Schalko et al., 2019; Muhawenimana et al., 2021).

We agree that including a summarised review of knowledge gaps and recent advancements would improve the manuscript. As such we have updated this section to provide greater understanding on the influence of large wood and LDs on geomorphology, as well as numerical and flume studies:

Despite their rapid deployment in riverine management over recent years, a key knowledge gap is how LD efficacy evolves temporally, both in response to geomorphic evolution up- and downstream of the LD, but also in response to flood sequences (Addy and Wilkinson, 2019; Grabowski et al., 2019). **The influence of large wood on river systems is well understood (1): wood increases fluvial complexity whilst being resistant to erosion and providing storage space for water (2) (Gurnell et al., 2018; Wohl et al., 2019). Specifically, large wood can form pools (e.g., Abbe & Montgomery, 1996; Al-Zawaidah et al., 2021; Ravazzolo et al., 2022), increase sediment storage (e.g., Comiti et al., 2008; Wohl & Beckman, 2014), protect against or induce bank erosion (e.g., Abbe et al., 2018; Galia et al., 2024) and influence floodplain morphology (e.g., Sear et al., 2010; Wohl, 2013). Large wood is generally mobile (Wohl et al., 2023), unlike LDs that are often engineered and anchored in-situ and therefore can be functionally different with wide-ranging designs (Lashford et al., 2022; Lo et al., 2022; Quinn et al., 2022).**

Challenges in disentangling the relative impact of LDs from the influence of land use, antecedent conditions and other flood risk management interventions presently result in an unclear understanding of their influence over time. **Similar to natural wood, LDs influence the hydraulic regime through increasing roughness and thus have the potential to influence channel geomorphology. The few empirical field studies that have focussed on LDs have**

highlighted that LDs can reduce peak flows for the 1-year annual exceedance probability (AEP) by 10% on average (3), however the response can be highly variable (Norbury et al., 2021; van Leeuwen et al., 2024). The backwater rise induced by LDs is also variable and can be increased or decreased with the presence of porosity-reducing material (Muhawenimana et al., 2023). Furthermore, the ability of a LD to store water, or sediment, can be dependent on the distance between the riverbed and the bottom of the LD, with gaps >0.3 m unable to store sediment in the Yorkshire Dales, UK (Lo et al., 2022), while increased wood volume also amplifies scour (Schalko et al., 2019). Laboratory experiments have shown that representing LDs as non-porous structures increases drag and flow area (Muhawenimana et al., 2021), and therefore it is important to account for porosity of the structures in numerical simulations. Yet often porosity is not considered in numerical simulations due to representing these complex structures in reduced-complexity models.

Recent works have focused on integrating LDs into 1D and 2D models at different spatial scales (Hill et al., 2023), most commonly representing the interventions as localised roughness adjustments (Pinto et al., 2019; **Geertseema et al., 2020**), geometry adjustments (Pearson, 2020; Walsh et al., 2020), or a combination of the two (Dixon et al., 2016; Senior et al., 2022) **(4)**. LDs have also been represented in hydraulic models, through stage-discharge relationships realising LDs (and other RAFs) as weirs or culverts (Thomas and Nisbet, 2012; Metcalfe et al., 2017; Keys et al., 2018; Hankin et al., 2019; Pinto et al., 2019; Hankin et al., 2020; Leakey et al., 2020; Pearson, 2020; **Follett and Hankin, 2022**). A comprehensive review of the large wood numerical modelling literature focused on artificially placed wood can be found in Addy and Wilkinson (2019).

Please see numbers inserted into text above:

(1) Assume this is a reference to naturally formed and occurring large wood?
(2) Do you mean large wood has resistance to being transported? Better state resistance to displacement instead to be clearer? The statement that large wood provides storage space for water seems a bit strange as by adding wood to a stream the volume of in channel storage would be reduced. Do you mean the ability of large wood to reconnect floodplains and create out of bank storage?
(3) Did both studies (Norbury and van Leeuwan) show the same reductions in peak flow for the same size of flow? Be clear.
(4) Are these hydraulic or hydrological models?

L64-65 Clear statements on potential geomorphic processes, feedbacks and importance are needed. At the moment the importance of considering geomorphic processes in models isn't coming through. For example, the aforementioned studies have given observations on the patterns of erosion and deposition in relation to structures that could have hydraulic effects.

Following on from the previous paragraphs that now more clearly state this, we have incorporated the following:

LDs and large wood clearly can alter local morphology, which in turn can alter hydraulic response through feedback cycles of erosion and deposition (Lo et al., 2021).

Ok.

L81-82 First part of this statement is not true. See references made above on progress made on representing the leakiness and lower gap effects of leaky barriers.

We have adjusted this sentence as follows to increase clarity around this point:

As such, no work currently exists that incorporates both the inherent 'leakiness' of LDs and the ability to simulate a lower gap **coupled with** sediment transport **to evaluate** geomorphic evolution **within a numerical model**.

Ok, this is better.

L82 What is meant by a 'prototype real world location'?

For clarity we have removed the word "prototype".

Ok.

L84 A new paragraph stating a clearer and more elaborated list of aims and, or hypotheses is needed. This would help to give the paper more structure and purpose.

We have added the following text in support of this:

The aims of this paper were to explore the relative behavioural impact of a simple LD on sediment transport processes and subsequent influences on discharge and water storage through a small reach. To do this, we first introduce new functionality for CAESAR-Lisflood that can represent LDs through the restriction of flow. Second, we evaluate the sensitivity of the model to DEM resolution, and third, assess the impact of LD roughness and gap size on geomorphology and water storage. Finally, we present the implications of numerically simulating LDs coupled with sediment transport processes to inform future modelling studies.

This is an improvement.

L159-160 'Upwinding' and 'upwind' are strange terms to use. Consider rewording?

The upwind scheme is the correct term for calculating the downstream flow field based on cells upstream, however for clarity we have updated this to:

CL employs a first order **upwind scheme (Coulthard et al., 2013).**

It would be better if this term was clearly defined here, or other terms were used to make this clearly understandable (or does this journal assume reader knowledge of the term?).

L190 This approach of scaling n seems sensible but perhaps a caveat is needed here given that it isn't based on an empirical relationship as it stands.

We have added the below caveat immediately following the introduction of BR:

BR captures increasing LD roughness with increasing stage to reflect increasing complexity as a greater vertical area of the channel is obstructed. Further empirical data is required to assess this assumption.

Ok this is an improvement.

L210-214 More details on the prototype reach is needed in the paper rather than citing the thesis. The channel slope used seems quite low for a headwater stream where leaky barriers are typically used. Would it be worth testing the model over a range of slopes to see the effect? Why were different DEM resolutions tested? Certainly important but little context or purpose on this is given.

We have added more details to this section as outlined below:

The model domain was 160 m long and 100 m wide **and represents a second-order stream**. The DEM had the same average slope as a prototype site (0.01 m m^{-1}; Wolstenholme, 2023) **where LDs were installed in 2019** and was created by linear interpolation between the high and low survey points in the reach captured with a Topcon OS-103 Total Station (TS).

We explored the impact of grid resolution to understand whether any bias was introduced in the model outputs, and to ensure that potential caveats as a result of resolution were highlighted in this development and technical paper. We agree that it would be interested to explore the impact of slopes on the effect, however this is beyond the current scope of the manuscript.

This helps but any details to give on the topographical survey in terms of survey strategy and point density? This can affect the quality of the topography used. What catchment area is the reach at? What is the channel width? Could testing different slopes be referred to as an aspect for further study at the end of the paper?

L221 What is meant by sediment types?

Sediment types referred to the different grain size bins used in CAESAR-Lisflood, however upon reflection this was confusing. Therefore, we have adjusted the sentence as follows:

This ensured that **the sediment was** distributed throughout the catchment in equilibrium with the topography.

Ok fine.

L238 Again like testing a range of slopes, it would be interesting to see the effects of a range of different flow events and perhaps would provide more insight into hydrogeomorphic effects of leaky barriers than testing a range of different DEM cell sizes.

As mentioned above, we agree that this would be interesting, however is beyond the scope of the development and technical paper where we focus more on the implementation of leaky dams in CAESAR-Lisflood.

Ok could this be referred to as an aspect for further study?

L246 Why was an nmax value of 0.16 used?

This was chosen as a conservative estimate of the leaky dam roughness based on empirical studies (e.g., Curran and Wohl, 2003; Dixon et al., 2016 and Addy and Wilkinson, 2019). Second, we explore the relative impact of nmax in section 4.3 from 0.12–0.2, and for the purposes of not biasing the impact of grid resolution or gap size with changing multiple parameters in parallel, we chose the central value (0.16) of this range.

To provide this justification in the text, we have made the following adjustments:
...dams was set to 0.16 **(chosen as a conservative estimate of LD roughness after Curran and Wohl, 2003; Dixon et al., 2016; Addy and Wilkinson, 2019)**.

Ok this is clearer.

Figure 6 To make it clear, mark on erosion and deposition labels. I.e. negative values show deposition and positive values show erosion which may be counter-intuitive at first glance. Similar remarks can be made for Figure 7 in relation to changes in water storage.

Negative values represent erosion and positive, deposition. We inverted the standard output from CAESAR-Lisflood to ensure that the elevation change was not misconstrued. We have, however, added text to that effect to Figure 6 to ensure that there is no confusion.
Figure 7 only shows water storage therefore we have not altered this figure.

Ok good, couldn't see the new revised figure to check this though.

L343-344 The finding that increasing LD gap size resulted in less sediment being lost seems counter-intuitive.

Less sediment was lost out of the model domain. There was less erosion upstream and downstream of the leaky dams when gap size was increased (e.g., Figure 9) as the river had less energy to erode the bed when compared to a smaller gap size.

Ok if that is the case but perhaps need to be clear that is for a condition of no sediment supply coming from upstream into the reach at the upstream boundary (if that was the case). I.e. are these geomorphic findings for a clear water scour condition? Adding a log jam in that blocks more of the channel for the latter condition would be expected to result in reduced sediment yield downstream.

L455-459 Perhaps a more important point is how valid is this model. Whilst there may be some value in using it in an heuristic manner for scenario testing, validation using empirical observations of changes in discharge and morphology would strengthen the value of this model and give practitioners more confidence in using it. This is a knowledge gap that should be stated. The word 'behaviouralist' seems a strange one to use.

We agree that to add more value the discussion would benefit from being more closely tied in with existing leaky dam literature. Therefore, to address this point we have included the following text at the end of section 5.2.:

Care must be taken when using the LD module for CL as the right results, such as elevation change and QQ reduction, may be overestimated when using coarser grid resolutions. **The scenarios simulated herein align well with the relative influence of LD from both field and laboratory observations including the formation of downstream pools (Lo et al., 2021; Lo et al., 2022; Muhawenimana et al., 2023), and the potential for sediment storage upstream**

(Comiti et al., 2008). Future work should focus on calibrating and developing this tool as a flexible and rapidly deployable option for LD simulations in CL, that should currently be used heuristically to mitigate the need for calibration. The LD module for CL can therefore be best used to understand the relative impact of LDs in larger, complex catchment to identify their individual impact on FRM.

This is better but the words **..'that should currently be used heuristically to mitigate the need for calibration'** are unclear and should probably be removed. The fact is the model is uncalibrated and unvalidated using real world data on morphological change adjacent to LDs. The potential value of the model to make predictions at the catchment scale is true but again some testing first at the reach scale is needed to underpin this first. This includes both calibration and validation.

L477-480 Yes this is a fair statement but there simply isn't usually the data to calibrate or validate sediment transport in the real world unfortunately and processes are notoriously difficult to predict for different hydrological events. In contrast, observations of channel hydraulics and hydrology are more available o accessible to enable model assessment.

We are glad that the reviewer agrees with this statement, and that the lack of data to calibrate/validate sediment transport highlights the need for the heuristic approach adopted and subsequent analysis.

Ok but model outputs in terms of erosion and deposition responses could be used for calibration and validation instead of sediment transport. Perhaps mention this earlier (see point above)?

L491 Could the model be used to assess other types of NFM structure or in channel wood? Another aspect deserving further study in future?

Absolutely. We have added the following to the previous paragraph to highlight this:

Additionally other types of structures with porosity can be evaluated with the model, if it can be defined through its height, gap (or lack of) and roughness, for example natural wood, bridges and bunds, and presents opportunities for further study.

Good though some of these structures are not necessarily porous (bridges, bunds).

Associated files with the paper. Appears that the model has been uploaded for both weblinks on Zenodo. Could not see the datasets.

All data used in the study and the model are available on Zenodo, in the .7z file.

Ok although the file format wasn't readable (CAESAR model needed?).

Typo errors

L467-468 This sentence does not make sense.

Apologies, this has been corrected to:

Sediment transport becomes increasingly important when unintended geomorphic adjustment to 'hard engineered' structures reduces the efficacy of structures, potentially increasing flooding downstream (Hesselink et al., 2003; Pinter et al., 2006; Hudson et al., 2008; Benito and Hudson, 2010).

Ok.

References

Follett, E. and Hankin, B., 2022. Investigation of effect of logjam series for varying channel and barrier physical properties using a sparse input data 1D network model. Environmental Modelling & Software, 158: 105543.

Geertsema, T.J., Torfs, P.J., Eekhout, J.P., Teuling, A.J. and Hoitink, A.J., 2020. Wood-induced backwater effects in lowland streams. River Research and Applications, 36(7): 1171-1182.

Lo, H.W., van Leeuwen, Z., Klaar, M., Woulds, C. and Smith, M., 2022. Geomorphic effects of natural flood management woody dams in upland streams. River Research and Applications, 38(10): 1787-1802.

Muhawenimana, V., Wilson, C.A., Nefjodova, J. and Cable, J., 2021. Flood attenuation hydraulics of channel-spanning leaky barriers. Journal of Hydrology, 596: 125731.

Norbury, M., Phillips, H., Macdonald, N., Brown, D., Boothroyd, R., Wilson, C., Quinn, P. and Shaw, D., 2021. Quantifying the hydrological implications of pre- and post-installation willowed engineered log jams in the Pennine Uplands, NW England. Journal of Hydrology, 603: 126855.

Schalko, I., Lageder, C., Schmocker, L., Weitbrecht, V. and Boes, R.M., 2019. Laboratory Flume Experiments on the Formation of Spanwise Large Wood Accumulations: Part II—Effect on local scour. Water Resources Research, 55(6): 4871-4885.

van Leeuwen, Z.R., Klaar, M.J., Smith, M.W. and Brown, L.E., 2024. Quantifying the natural flood management potential of leaky dams in upland catchments, Part II: Leaky dam impacts on flood peak magnitude. Journal of Hydrology, 628: 130449.