In this study a landscape evolution model, CAESAR-Lisflood (CL), is applied to a steep mountain catchment to assess the effectiveness of engineering works in reducing the transport of sediment. This is an applied study, that is straightforward, and demonstrates the use of CL in a highly dynamic landscape. Overall, the manuscript fits the scope of NHESS and would be interesting to modellers and practitioners working in mitigating geo hazards in mountainous regions. My concerns with the study are related to the choice of hydrological parameters, the physical plausibility of landscape changes, and development of initial conditions. In addition, the clarity of the manuscript requires substantial improvement and I recommend the text is thoroughly edited by a native English speaker, with a background in fluvial geomorphology, before acceptance. Below are major comments that need to be addressed followed by a list of minor points and edits.

Major comments

A weakness of the study is the lack of calibration of the hydrological component in CL. As such, there is no way of knowing if the quantity and timing of the floods in the ungauged catchment are accurately replicated by CL. The hydrological parameters adopted (m-values) are from studies performed from nearby catchments but these studies also have not performed calibration to derive m-values. The authors, instead rely on landcover to assign m-values, but m-value is only partly dependent on landcover. For example, Ramirez et al.2022 found that in a mountain catchment soil depth correlated well with m-value and not with landcover. To have greater confidence in the model, the authors need to provide hydrographs for the entire simulated period and, in addition, provide qualitative or quantitative data that confirms the physical plausibility of the simulated discharge, specifically the floods.

In this study, CL simulations have produced locations of deep erosion between 3-10 m in a period of three years. This is quite a bit of erosion in such a short period and in some instances would produce features in the simulated landscape that resemble small canyons. Could you verify that these erosional features are physically plausible by providing photographic evidence from the observed landscape and compare them to cross-sections from the simulation. Or provide any other type of validation that supports such extreme erosion across the simulated landscape. In addition, across all simulations (Fig. 5a), there are instances of erosion that exceed 3 m in the downstream area where erodible thickness is 3 m. Can the authors explain how simulated erosion can exceed the thickness of the initial erodible sediment? Likewise, in this study, how is it possible for CL to produce erosion between 10 and 15 m, if the maximum depth of erodible sediment in the catchment is 10 m?

In the study there is no mention of establishing initial conditions by spinning-up the model to mix the grain sizes. If spin-up was not performed, can the authors provide an explanation. If spin-up was performed, could you briefly explain how it was done in the methods. Regarding the choice of bedrock elevation (Fig. 2), could the authors provide the physical basis for the choice of erodible thickness values and locations of these values.

Modelling the long-term geomorphic response to check dam failures in an alpine channel with CAESAR-Lisflood. International Journal of Sediment Research, 37(5), 687–700. https://doi.org/10.1016/j.ijsrc.2022.04.005

Minor points and edits

Line 20: replace Take with Taking

Line 29-30: “Moreover, the effectiveness reduced gradually caused by the storage capacity of dams decreased.” needs rewording.

Line 46: be more specific than just stating “complex processes and origins”

Line 47: “treatment” doesn’t sound right

Line 50: “without more practices” doesn’t sound right

Line 52: “The subjective expression” has not been defined, so most readers won’t know what this means

Line 56-57: define “long time effects” and “short-medium term” is it 10s, 100s, 1000s years?

Line 59: what is meant by “special sites”?

Line 60: add reference for CL.

Line 61: Cellular Automata

Line 67: add Ramirez et al. 2020 and Peleg et al 2021 as recent examples of simulating channel evolution with CL

Ramirez, J. A., Zischg, A. P., Schürmann, S., Zimmermann, M., Weingartner, R., Coulthard, T., & Keiler, M. (2020). Modeling the geomorphic response to early river engineering works using CAESAR-Lisflood. Anthropocene, 32, 100266. https://doi.org/10.1016/j.ancene.2020.100266

Peleg, N., Skinner, C., Ramirez, J. A., & Molnar, P. (2021). Rainfall spatial-heterogeneity accelerates landscape evolution processes. Geomorphology, 390, 107863. https://doi.org/10.1016/j.geomorph.2021.107863

Line 68: add Ramirez et al. 2022 as another recent example for applying CL to dams

Modelling the long-term geomorphic response to check dam failures in an alpine channel with CAESAR-Lisflood. International Journal of Sediment Research, 37(5), 687–700. https://doi.org/10.1016/j.ijsrc.2022.04.005

Line 75-77: “assess the effectiveness of a set of mitigation facilities” in doing what? In reducing channel change, sediment transport…be more specific.

Line 98-99: “The earthquake strengthened the solid material produced and reached 10^6 m^3” needs rewording

Figure 1: Remove contours to improve clarity of map (besides contours are not labelled and have limited use). “Translation area” should be replaced with “transitional area”, and the same should be done throughout the entire manuscript. Dams are barely visible in map, perhaps fill them with a color. Spell out “Figure” instead of using the abbreviation and do this for all captions.

Line 107: replace “motivated” with “transported”

Line 108-109: “here burst 6 group debris flow-flash flood disaster chains in rainfall season according to field survey” needs to be reworded and 6 needs to be spelled out.

Line 110: replace “occurred time” with “time of occurrence”

Line 111: replace “remote sensing image data” with “remotely sensed image”. In addition, provide more information about the image, like the name of the satellite, spatial resolution and date of acquisition. Also, provide more information on how you determined the location of the landslides.

Table 1: “mess of farmland” needs to be reworded. Maps in table are quite small, perhaps remove them from the table and make them larger in a multi panel supplemental information figure, and

reference this in the text. In the table you could add the summary statistics of the landslides derived from the maps (e.g. total area, min area, max area).

Line 127-128: replace with “The upper dam has storage capacity of x m^3 and a height of m, and the transitional area dam has a storage capacity of x m^3 and a height of m”

Line 128: replace “With deposited in the reservoirs gradually,” with “With the reservoirs gradually filling with deposits,”

Line 130: what do you mean by “lower reservoir was full of loose material.”, be more specific.

Line 133: quake-stricken

Line 134: “Control processing” needs rewording

Line 145-146: Provide an example of a vegetation revetment.

Line 147: replace “plow” with “agricultural”

Line 152: Change to NHESS citation format: (Tom J Coulthard et al., 2013)

Line 152: Change to NHESS citation format: (T J Coulthard et al.)

Line 155-158: This text needs rewording

Line 161: “Besides the creative flow model” needs rewording

Line 163: “slope progress” needs rewording

Line 166: CL has sediment transport equations, provide the use of these function in the model description.

Line 170-172: This text needs rewording

Line 173: “we reconstructed four parameters” is not entirely correct wording because the DEM and bedrock layer are model initial conditions, the rainfall is a model driver, and the m-value is a parameter.

Line 183-184: “which were prone to form by interpolation operation, and then caused the hydrological module to calculate inconsequently.” needs rewording

Line 185: Explicitly mention if the UP DEM does or does not contain dams.

Line 186-188: should be reworded to: “the present-day protected landscape surface DEM (PP DEM) included the dams by raising the grid cell elevations by 10 m for the upper dam and 9 m for the dam in the transitional area”

Line 189: replace “extracted” with “produced and “

Line 188-191: should be replaced with “The enhanced protected landscape surface DEM (EP DEM) includes the dams in PP DEM and, in addition, levees were represented at selected locations. Levees were produced by raising grid cell elevations by 2 m.”

Line 192: You mention a field survey and how this led you to conclude the thickness of erodible sediment was spatially variable. In the text, elaborate more about the field survey and how it informed the production of the bedrock layers in CL.

Line 194: replace “different” with “heterogeneous” or “variable”

Line 194-195: needs rewording

Line 200: remove “supposed to be”

Line 200-204: should be “For the river channel and outlet, there would be a large amount of deposition and the thickness of erodible sediment was set to 5 m and 4 m respectively. The dams in Scenario PP and dams and levees in Scenario EP were non-erosive concrete. As such, the erodible thickness of these features was set to 0 m”

Line 205: define bedDEMs and “In addition, DEMs were formatted to ASCII raster as required by C-L”

Fig 2.: The flowchart is redundant as the text has already described the process quite well. I like the figure, but without the flowchart and then change the caption accordingly.

Line 213-221: should be like this: “Another parameter set in each scenario was the m value in CL’s hydrological model (TOPMODEL), which controls the exponential decline of transmissivity with depth (Batty et al., 1997) and influences the peak and duration of the hydrograph in response to rainfall. The lower the m value, the lower the vegetation coverage, the flashier flood peaks, and shorter flood hydrograph duration (Citation needed). In this research, the m value in the UP and PP scenarios were set to 0.008 without spatial variation, which represents the vegetation coverage of farmland as determined by Li et al., (2020) for a catchment nearby with similar landcover. As mentioned earlier, the upstream-low elevation area covered by the biological measures designed in the EP scenario was assigned a higher m value.”

Replace all instances of “m” with m value throughout the manuscript.

Figure 3: Capitalize the y-axis label. Caption should be like this “Fig. 3(a) Daily precipitation in 2011-2013 (the red vertical line indicates daily maximum precipitation of 126.5 mm); (b) Hourly precipitation in 2016; (c) Downscaled hourly precipitation in 2011-2013 (the red horizontal line indicates the hourly-mean precipitation 5.27 mm in the day with maximum precipitation showed in (a)).”

Line 227: change “(Fig. 3(a))” to (Fig. 3a), and do the same throughout the manuscript for all references to multi panel figures.

Line 229: Change to NHESS citation format: (Tom J Coulthard et al., 2012)

Line 243-244: I am not sure what you mean by “the downscaled hourly precipitation series was better than the hourly mean precipitation”, please reword this sentence.

Table 2: Explain the choice of setting the slope failure threshold so high (60 deg), wouldn’t this prevent the occurrence of landslides and lead to deep canyonlike incision?

Line 263: remove “internal” and any use afterwards because internal geomorphology does not mean anything

Line 264: should read “assess the effectiveness of the interventions.”

Line 265-266: should read “The simulated annual landscapes were analysed to quantify the geomorphic change, and were derived…”

Line 269-270: Here you mention “damage” but what you are really calculating is exposure because you are not calculating a monetary value. Change the text to mention exposure, and I am assuming that a map of settlements or landcover was used to calculate the exposure, if so, provide this map in the supplemental information. If you didn’t use a landcover map, explain what you used to derive exposure.

Line 274: remove “internal”

Line 284: explain what a low and high value means for CA

Line 292: Is this daily sediment yield measured at the catchment outlet? If so, change the text accordingly.

Line 303: replace “panoramas” with “landscapes” and do the same throughout the entire manuscript.

Figure 5: Hillshade does not need a legend. Add quantitative values to the erosion and deposition legend. Place one scale bar in a map. In panel (a), remove all dates in maps and place dates once above the top row of maps. In panel (b), it still isn’t clear how damage was calculated, but see my comment above. I think caption (c) should read “the distribution of deposition and erosion at the conclusion of the simulation for the three scenarios.”

Line 308: “aggravated” is not the right word

Line 310-311: Qinggangping gully and Shicaozi gully need to be labelled once in Fig. 5 because the reader is not familiar with these locations.

Line 312-315: Again, this is exposure not damage, change the text accordingly.

Line 317-327: The narrative you provide is fine, but you need to provide some values that support your comparisons.

Line 322: “internal geography” does not make any sense

Line 333-334: should be “Here, we provide a detailed investigation of the controlling measures and surroundings for the three scenarios.”

Figure 6: Unbuilt dam is not legible, fill with a color. Built dam needs to be a different color than red to provide contrast with areas of deposition. No biological and biological protection fill patterns appear identical and make it difficult to see the landscape changes. Just have no fill pattern. The erosion and deposition legend needs a title like “Geomorphic change (m)” and in the caption mention that negative geomorphic change values are erosion, and positive values are deposition. Add an inset map that clearly defines where the key spots are in the catchment. Add a scale bar(s). Label dam 1 and dam 2. Caption should be “Figure 6. Geomorphic changes at the conclusion of the simulation at key spots for the UP, PP, and EP scenarios. Top row is the upriver section containing dam 1, dam 2 and the vegetation revetment. The bottom row is the downriver section containing levees.

Line 341-351: Add values to support the existing qualitative comparison.

Figure 7. Capitalize the y-axis label. Mention in the caption that scenario UP does not have dams, but deposit depths are provided where dams existed in the other scenarios.

In Figure 6, the EP map shows the levees blocking a tributary, is this a mistake in the figure or did you really block this tributary in the simulation. Please explain.

Line 372: “Divisional erosion and deposition” doesn’t make sense

Line 373-379: Add values to support the existing qualitative comparison.

Figure 8: I am unsure what is plotted here or what type of plot this is. Is this the distribution of simulated geomorphic change after three years in the source, translational, and deposit areas? If so, are these boxplots? If these are boxplots, provide the median, IQR, whiskers and outliers. Add more explanation to the figure caption and the y-axis label should read “Volume of sediment”.

Figure 9: caption should read “(cyan shading represents when PP is more effective than EP and red shading represents the opposite)”. Be consistent when labelling panels, as here you use a) and in previous figures you use (a). Do this for all of the figures.

Line 408: replace “What’s more” with “Additionally”

Line 410: “indeterminately” doesn’t make sense

Line 464: Briefly mention the method applied by Ramirez et al. 2022 for check dam failure.

> Modelling the long-term geomorphic response to check dam failures in an alpine channel with CAESAR-Lisflood. International Journal of Sediment Research, 37(5), 687–700. https://doi.org/10.1016/j.ijsrc.2022.04.005

Line 512: “debris shocking” doesn’t mean anything

Line 513-514: “In addition, the decrement in EP suggested the accumulated materials blocked by dams upgrade a slope upstream in turn.” needs rewording

Conclusion section: Here you need to further summarize your main findings because the current text reads like a repetition of the results.