

Comments on manuscript, 'egosphere-2024-1191'

General feedback

First, I have noticed that another reviewer has provided feedback and the authors have updated their manuscript to reflect those comments. **My thoughts below are therefore based on the revised manuscript uploaded on the 6th July 2024.**

The authors present an important and potentially powerful tool to aid in the calibration of complex numerical models. CAESAR-Lisflood, as a model that takes several parameters, many of which can be sensitive to small changes in values, is an ideal model to develop the iterative model calibration (IMC) framework. The conclusion section outlines very effectively how this IMC framework can be adapted for other use cases and in my view, persuasively argues how this is not just confined to CAESAR-Lisflood or modelling gully erosion.

I outline a few core questions/concerns, and list several minor corrections that I recommend the authors should address:

- Why did you choose gully erosion in Australia as your case study to demonstrate your new IMC? I don't doubt there are plenty of good reasons, which may include your familiarity with the environment and/or processes in question, but it wasn't clear to me as a reader why this specific case study is important. Could you elaborate on this in the "Problem statement" section?
- You mention in section 4.1 "Problem statement": simulating at a range of different temporal resolutions (e.g. days, weeks, months – as well as annually). Yet, in terms of observational data, you essentially have just 1 datapoint to compare against – i.e. the net erosion volume and gully morphology differences between the years 2019 & 2021, with no other observations in between. Given the observational data you present, I don't think you can argue that you can accurately simulate changes at fine temporal resolutions. For instance, gully erosion over the 2019-21 time period could be driven by hundreds of small events; a smaller number of intermediate sized events; or 1-2 very large events. You won't know which of these scenarios will apply if you've only got DEMs for 2019 and 2021. I think it would be more appropriate to avoid the subject of modelling at multiple finer temporal resolutions of days, weeks, months, etc.
- You also mention extrapolating beyond 2019-21 and argue that modelled geomorphic changes are "justifiable" and "follow a consistent trend over time". Yet, you don't have the data to back these claims up. I would advise then that you either: a) remove this from the paper entirely, or b) compare your model results to more recent post-2021 observed geomorphic change data – assuming this exists of course!
- You write about "calibration epochs" and describe these as a function of "rounds x iterations". Yet you don't explain clearly what "rounds" and "iterations" mean here. Without this explanation, it's very difficult to interpret the results and discussion you present in section 5.2 "Experimental analysis".
- I don't think you have explained the importance of the experiments to evaluate IMC's efficiency in CL parameter re-estimation. Why does it matter? And are there single optimum values for parameters like "lateral erosion rate"? Perhaps I'm missing something obvious here...

Specific issues

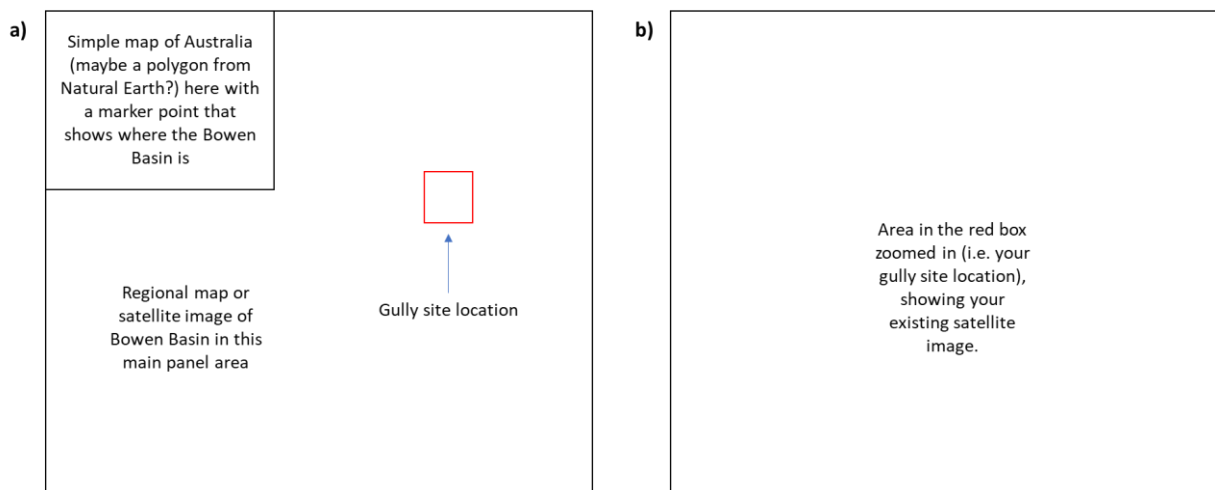
Text font on most figures is very small. I have to zoom in quite a lot to see the text on my computer screen and I am certain the text would be unreadable on a printed copy. Please could you make the following amendments?

Figure 2: You should move this into section 3. Currently, the figure is introduced at the end of Figure 2, but crucial details like the “prior data (PD) file” get explained in later paragraphs; this current order of presentation is confusing to read.

Also, “CAESAR-Lisflood” is spelt incorrectly on the figure itself.

Figure 3: I would consider starting again from scratch with this figure. How about something like this:

Panel (a): A regional map that shows the Bowen Basin of Northern Queensland, with an outline that shows where your specific gully site is located. You could include here an inset map in the top left corner that shows Australia, with a shape or marker point that indicates where Bowen Basin is located. Then panel (b) would show your existing satellite image of the gully site zoomed in – see sketch below that shows what I mean. Below panels (a) and (b), you could have panel (c) showing your pluviometer reading.



You already include the DEMs as part of Figure 4, so I don't think it makes much sense including them in Figure 3. I also think you can give a clearer introduction to the context of your study site to the reader by re-formatting the figure to how I've suggested.

Also, please make sure that the font size of all text on the figure itself is larger than it is currently; it's very difficult to read even when zooming in on a computer screen! You may also want to put the satellite image credit text in the figure caption itself to allow readers to see it properly.

Figure 4: Same comments on font size apply here as they do for Figure 3: the numbers on the axes of the graphs depicting erosion volumes need to be enlarged. You may find that this figure needs to be displayed in landscape format to show everything properly.

Figure 5: Expand the figure caption slightly to explain what you mean by “standardized standard deviations”. You haven't explained what this means anywhere in the main text; adding it to the figure caption would help readers to more quickly interpret this figure at a glance.

Figure 6: Please increase the font size of labels on the axes and the vertical colour bar legends.

Figure 8: Please increase the font size of labels on axes and the 2 category labels on both plots (i.e. the yellow and orange bars). It may make sense to put part (b) below part (a) to help make the figure more legible overall.

Some parts of the text require further clarification:

In several places in the results, you use the word “Actual” to refer to the observed gully changes. Why not just call it “Observed” instead for clarity?

Lines 56-59: You mention that Tsai et al. (2021) propose a couple of data-driven approaches for calibration, yet you only describe one of these. What is the other approach?

Line 74: Remove the first word of the sentence “Besides a large number of...”

Lines 78-79: I would remove the sentence: “Our calibration approach innovatively leverages...”. At this point in the text, you haven’t yet demonstrated that this is the case, and it reads more like a sentence for the conclusion section of your paper.

Lines 92-101: I don’t know if you need this explanation of what each section specifically covers. Perhaps replace with a couple of sentences explaining that you introduce this new IMC algorithm, demonstrate how it works for a chosen landscape evolution modelling context (i.e. gully erosion modelling with C-L), and outline how this IMC could be adapted to other contexts going forward.

Line 131: Replace “it’s” with “which is”.

Line 155: Should “Numerical” begin with a capital letter?

Line 203: Citation should be written as (Nash and Sutcliffe, 1970).

Line 219: You write “...we introduce the study area...”, yet you’ve already introduced the study site in the preceding paragraph. Perhaps delete or move this sentence?

Line 265: What does “w.r.t.” mean here?

Line 273: Surely the “target erosion volume” is derived from the difference between the 2021 DEM and the 2019 DEM, not simply the 2021 DEM itself? Rephrase.

Table 3 caption: You mention apparent high variability in parameters 10, 11 & 12. How do you arrive at that judgement? Did you look at the coefficient of variation for all 12 parameters? It might be worth including this as an extra column in your table.

Line 274: “We provide a visual comparison of the same in Fig. 4...” Same what exactly?

Section 4.3.3 Future land evolution: Demonstrative projections: I don’t think you have address Reviewer 1’s comments adequately here:

- How have you run future projections here – by repeating the 2019-21 rainfall file?
- How do projected future landscape changes follow a consistent trend over time? Are you suggesting here that the annual rate of erosion does not change compared to the 2019-21 calibration period?
- I don’t understand what you mean here by “justifiable landscape changes”? I know you’re referring to erosion volumes, but what exactly makes them justifiable here? I agree with

Reviewer 1 that you would likely need to look at more recent observational evidence post-2021 to back up what you're trying to say here.

Line 291: What do you mean by “fixed type of numerical model”?

Line 300: What do you mean by “rounds” and “iterations”? These terms could easily be used synonymously, so it would be very helpful if you could explain what you mean by each of these.

Lines 306-309: I would consider merging with the preceding paragraph. It took me a couple of minutes to work out what you meant by “This phenomenon...”

Lines 349-351: You note an important caveat to using MSE as an evaluation metric for your modelling. Is it worth adding a sentence here to mention that your IMC framework could easily be adapted to use an alternative metric instead, should the user feel that that is more appropriate?