Title : Phenomenological model of suspended sediment transport in a small tropical catchment

General comments

The authors present a new model to simulate suspended sediment concentration based on water depth time series, for floods that exhibit a counterclockwise hysteresis pattern between concentration and discharge. After calibration on the Capesterre tropical catchment, the model appears to be able to simulate correctly such hysteresis patterns, and the concentration values. Moreover, the calibrated parameter values are consistent with physical explanations of river processes. The model is also tested against another catchment where it provides consistent results.

The paper addresses the issue of suspended sediment modeling, which is clearly relevant to the scope of SURF. The proposed modeling approach is novel, yield convincing results, and will certainly of great interest for the community. The paper is clearly written. My main concern is that I find this paper is not well rooted into the literature, particularly concerning suspended sediment modeling. This should be improved both for the introduction and discussion sections (see details below) in order to make the paper stronger. Moreover there are several points that deserve more accurate explanations and discussion (the concentration-turbidity relation, the model assumptions, the quantification of model performance, see line by line comments) and some conclusions that are a bit too quick. I therefore recommend major revisions before accepting this article in ESURF journal.

Introduction: I find that the authors do not relate their work to existing literature on suspended sediment concentration or yield models. There is a large body of research concerning the modeling of suspended sediment yield, with either advection-settling-erosion equations, concentration-discharge relations, empirical models at various time-scales …. Therefore, it would be useful for the reader to provide some background on existing approaches, explain their limitations and show how the proposed approach brings novelty and improvements. I am totally convinced that it does bring novelty but this has to be explained in relation to the literature.

Discussion: Currently, the main discussion is about the meaning of the parameters (section 5). Even if this is a very interesting section, I think that the paper deserves a larger discussion section, including (3) a discussion of what this new model brings in comparison to existing models of suspended sediment yield, (4) a discussion on the limitations of this model, in particular concerning the validity of the assumptions, and the uncertainty related to the concentration-turbidity relation, and (5) some perspectives and applications of such model, for instance at larger time and spatial scales. I suggest adding a full discussion section to the paper, which could incorporate the existing sections 5 (meaning of parameters) and 6 (test against another data set), and address all the other points mentioned above.

Structure: I might be a bit conservative but I am not totally convinced by the current structure of the paper. It reads well but I find that it does not make clearly visible what the results of the paper are. Is the field data from Capesterre a result from this paper, or a preliminary summary of previous studies, only presented here as an introduction for the modeling work? I would suggest using a more standard organization with a methodological section (field setting, field data acquisition, model construction and calibration), a result section (field data sets and calibration results) and a full discussion section that could address all the points mentioned earlier. This is not a critical point for publication however and I would let the editor comment on the necessity, or not, to change this structure.

Line by line comments

Line 2: “the model correctly represent the transport of suspended sediment”: I think this sentence is not exact; the model actually does not represent the transport of sediment itself. I suggest replacing by “the model correctly represents the evolution of sediment concentration at the outlet”

Line 6: “both of which are related to the availability of fine sediment”: I don’t think that this is demonstrated in the paper. I suggest removing this from the abstract.

Line 47: Perhaps worth mentioning that the distance of sediment sources to the outlet is an important factor of hysteresis. And that this effect can be coupled with the spatial heterogeneity of rainfall.

Line 51: Here I think that the authors should review existing approaches for modeling suspended sediment transport. There had been a lot of research on this topic so they should at least explain what the limitations of existing approaches are and why they choose to build a new type of model.

Line 94 and figure 1d: The calibration of the turbidity-concentration relationship is quite noisy. Could the author introduce an estimation of the uncertainty for the concentration extracted from this relation?

Moreover, is there any information available on the variability of suspended sediment grain-size during the floods? It is sometime argued that hysteretic patterns in concentration may result from changes in the turbidity-concentration relationship due to grain-size variability (Landers and Sturm, 2013, https://doi.org/10.1002/wrcr.20394). Figure 1d suggests that there is a great variability in this concentration-discharge relationship, which could be related to a variability in grain-size. **This is an important point to check before attributing the hysteresis in turbidity to a hysteresis in concentration, which is the focus on the paper.**

Update: having read the whole paper and seen figure 6, it seems that there is indeed a shift in grain-size during the flood. I suggest that the authors use this data to estimate a mean grain-size on the rising and falling limb, then search for information from the sensor manufacturer, and/or from the literature, to estimate if this grain-size shift could affect, and how much, the turbidity-concentration relation of Figure 1d.

Line 95: Is the field data on water level and concentration already published somewhere, or could it be published with this paper? I find it a little frustrating to read about a 3 year data set with 217 floods without seeing more than what is displayed in Figure 1e, 2 and 4. Even if I trust the authors concerning their analysis on flood extraction and hysteresis calculations, I think it would be a good thing to make this data set available, both for the reader, and for the data producers themselves.

Figure 2: The three floods that are selected here have a very simple shape, with an abrupt increase and an “exponential-like” decline. Are these floods representative of the whole data set, do most floods exhibit the same shape in this catchment? Are there any multipeak events?

Line 145-147: Do you think that this bed replenishment from hillslope really occur gradually, with no effect during floods? I would expect that hillslope processes such as landslides or debrisflows should be more frequent during strong rainfall, therefore could also represent sediment sources during floods on top of the riverbed source itself.

Line 148: If available, an information on suspended sediment grain-size would be useful here. If it is mostly silt, rather than sand, it would reinforce this assumption

Line 152-155: **This is indeed a very strong assumption, which is not realistic at all to me. It raises the issue of water mass conservation**. With a uniform rain, the amount of water arriving at one point should scale with drainage area. Therefore, it is not possible for the water depth to be always the same at every point in the system. Even if you argue that the river width increases downstream, this increases is generally thought to scale with drainage area to a power < 0.5 (see Lague 2014, DiBiase and Whipple, 2011 for instance) therefore this increase is not sufficient to conserve water unless water depth increases with drainage area too.

To my mind, this can not be presented as an “simplifying assumption driven by field observation” (line 140) since it does not seem to be rooted in any observations, physical principles, or expected behavior of the system.

Perhaps an alternative approach to simplify the model would be to write the full equation, then make the assumption that the advection terms are negligible (with no a priori physical basis), then check that this is indeed the case after solving the model. This could be done by comparing estimates of advective fluxes, and erosion and deposition rates after calibration.

Line 156: Removing the advection terms: does this imply that there is no flow velocity at all (which sounds problematic since bottom shear stress (induced from flow velocity) is responsible for sediment entrainment? Or is it that flow velocity and concentration are the same upstream and downstream therefore the advection contribution is null ? I understand that it should be the second option, but this should be made clearer to the reader.

Line 167: Perhaps you could detail that an increase of h implies an increase of the flow velocity hence of bottom shear stress?

Line 168: This form of erosion rate is not new, I think it is a classical form for erosion rate as a function of a bottom shear stress in the sediment transport literature. The author could acknowledge here that they build the model based on these existing approaches.

Line 204 and further: Why using “assimilation” instead of “calibration”? As far as I understand, the procedure described here is a model calibration?

Line 209-211: “the model reproduces surprisingly well” and “the model accounts reasonably well”: **could you quantify how well does the model perform, using a quantitative index**? For instance, a RMSE, or other. This would be useful for comparison with the results over a longer period, and the results on the Laval catchment.

Line 215: Any suggestion to address this bias? Perhaps this could be discussed in a future discussion section?

Line 225: It would be interesting to compare the model results with the results from a simple rating curve approach (fitting a relation between Q and C), both at the scale of the event, and of the chronicle.

Line 230: is it possible to run the same test over a full year of data ?

Line 233: “reasonably well”: can you quantify this?

Line 240: Would a model based on a simple rating curve perform as well, or better, over this period? The present model definitely brings some improvement in terms of simulating the hysteresis, but does it also improve the simulation of total sediment export compared to a simpler model? This could be quantified by calibrating a simple power law relation between discharge and concentration, calibrating this model against the same data and computing the same performance index (RMSE or other) than for the present model.

Line 248: “we suspect that the characteristic erosion rate and the exponent n reflect changes in the sediment availability”: **Could you explain a bit more? Why? Any suggestion to test this hypothesis ?**

Line 261-270 : Here you use an expression for the drag force (equation 7) which is correct for a turbulent flow, but the particle Reynolds number (Rep = Vs ds/nu) is smaller than 1. Therefore you should rather use the classical Stokes formula, Vs = (s-1)gD^2/(18nu).

(And I don’t think that Andreotti et al (2013) were the firsts to introduce Stokes’ law by the way ☺)

Line 290: Do you mean that the water level is a proxy of the bed shear stress?

Line 297: This equation (12) requires an assumption of flow uniformity, which is worth mentioning to the reader.

Line 301: What is the resolution of the DEM that is used to compute this slope? It is usually quite difficult to obtain accurate river slope values from a DEM unless the resolution is really fine. Moreoever, using equation 12 to estimate the shear stress requires that the flow is uniform at the gauging station, which is often not true (when stations are located on a weir or in a narrowing section for instance). Is it (approximately) the case at La Digue station?

Line 310-317: This is an interesting point to discuss, particularly in relation to Misset el al, GRL, 2021, that showed how bed mobility was able to release fine particles and increase concentration. **However, this result is in contradiction with the first assumption of the model (Line 144) according to which there was a large amount of fine sediment available in the riverbed.**

I understand that for a large flood, coarse bed sediment can be entrained therefore release fine sediment from the subsurface. But there should be a range of intermediate floods for which the coarse particles do not move. For such floods, either there is a large amount of fine sediment on the bed hence the critical shear stress should correspond to a small sediment size, either fine sediments are mostly stuck in the subsurface hence the transport-limited assumption does not hold. This should be discussed into more detail. Perhaps you could calibrate the model on floods of various amplitude and check if there is a different trend in the calibration for intermediate and high floods? Otherwise, I think it is a bit too early to conclude in a general way from these 4 values of ht that “the threshold of suspended sediment transport is set by that of the coarse particles”

Line 399: “the model represents rather well”: What do you mean by “rather well”? Could you quantify? (see earlier comment)

Line 341: “the best-fit parameter values fall within the same ranges as those obtained in Capesterre”. **So what does it mean? Could you discuss on the potential implications of this observation?** Does a similar range of falling velocity suggests that grain-sizes are the same? Does a similar range of threshold water level indicate that suspended sediment transport is also controlled by armoring in the Laval catchment?

Line 347: “that accounts for the transport of suspended sediment”: I don’t think that the model accounts for the transport (no advective terms). I suggest replacing by “that represents the temporal evolution of the concentration during floods.

Line 349: same comment as above

Line 356: This has only been suggested in line 248 but has not been discussed or demonstrated. I don’t think that this should be part of the conclusions. “We suspect that” without any other kind of explanation is not a scientific result in my opinion.

Line 365-368: **I think these points deserve to be addressed in a discussion section**. In particular, a discussion concerning the limitations of the model, and the effects of grain-size, should be developed in the paper earlier than in the conclusions.