1 Responses to reviewer 1

I have carefully reread the new manuscript and the authors' response to reviewers. My original comments about the scientific value of this work stand, and I am pleased to see that the authors have greatly improved and expanded the introduction and discussion to better contextualize this work and address salient questions. I have a few minor comments but I don't think another round of review is necessary; this is a valuable work worthy of publication.
We thank you for your appreciation of the revised manuscript. We detail below how we addressed your additional comments.

L92: "before concluding" – you could tease in short here what the most important conclusions of this study are.
We have modified part of the last paragraph of the introduction to tease the most robust conclusions of our study.

L134: is this a way of notating the uncertainty in that measurement? Maybe add + to the superscript and – to the subscript to make that more clear.
This notation is not a way of notating uncertainties, but a way of indicating a range of variations, sometimes used in planetology. We understand that this is confusing. Accordingly, we have rephrased the text which now reads "It is virtually zero 98% of the time, and rises only during floods, where it may reach up to 1430 mg L^{-1}."

L145, fig 2: I still think a dashed vertical line in the left subpanels aligned with the flood peak would help to visualize the time lag of the sediment peak.
This is done for fig. 2 and Fig. 3, thank you for the suggestion.

L153: "ignoring the washload" – can this be justified in some way? How much does the washload concentration change much during the flood? My concern is that turbidity is often extra sensitive to the finest sediment.
We agree and have modified the text accordingly. In fact, the data at our disposal show that the distribution of sediment size is bimodal, with two peaks. The first one, around 5 micrometers, corresponds to the washload, and remains remarkably stable during the flood. The position of the second peak varies between 23 and 32 micrometer for the data at our disposal. Assuming that these data are representative, we expect no significant change of the turbidity-concentration relationship during a flood. We have modified the text to clarify this point.

Fig 3: I know it's the same as fig 2 but nevertheless the figure should have a legend. Fig 2, Fig 3: Add null model (c=f(h)) as you did with fig 5?
We have added a legend on Figure 3, as requested. However, the addition of the null model to Figures 2 and 3 makes these figures cumbersome, without providing any new information : these figures display loops that cannot be represented by the null model. For this reason, we prefer to restrict the discussion of the null model to the time series of fig. 5, only.

L173-174, Fig 4: consider adding as a subpanel a cartoon of how the hysteresis index works. For me equation 1 was not obvious, but a visual example of the integral makes it immediately obvious why it works.
We added a subpanel on both sides of the H_I graphically explaining the calculation.

L257-276 very clear explanation!
Thank you for your appreciation.

L279: suggest omitting "bed" from "bed area" as this makes me think you're talking about concentration in the bed rather than in the flow.
Done.

L308, L311: I think "calibration" would be a better word than "adjustment"

We have rephrased this sentence which now reads "In the next section, we discuss the calibration of our model on field data, over a short period of time corresponding to a single flood event.

L446: "there, the flow is approximately uniform" – what is this based on?
It is an assumption. We have clarified this point by rephrasing the sentence which now reads : "There, assuming that the flow is uniform, we approximate the threshold shear-stress by ...

L505: focusing
Corrected.

L530: why would the simple rating curve systematically overestimate? This seems like a problem with the fit; I would think some floods would be overestimated and some underestimated, with higher RMSE than your method, but not necessarily a systematic difference.
We constructed the rating curve from direct measurements of the concentration of suspended sediment in water samples collected by an automatic sampler all through the year. The automatic sampler is equipped with a pressure probe, which triggers the collection of 24 water samples, one every 15 to 60 minutes, as soon as the rivers stage exceeds a threshold set by the operator. In the dataset at our disposal, this threshold varies between 30 and 80 cm, that is way above the threshold of entrainment in the Capesterre river. Because of this, most measurements belong to the falling rather than to the rising part of the hysteresis. This biais likely explains why the rating curve overestimates the concentration of suspended sediment in the river. We have modified the corresponding paragraph to clarify this point.

2 Responses to reviewer 2

First, I congratulate the authors for the improvements made to this revised version and particularly concerning the introduction and discussion revisions. I also thank the authors for the detailed and convincing answers to my questions.
We thank you for your appreciation of the revised manuscript. We detail below how we addressed your additional comments.

Line 153: Why should you ignore the washload when considering the possible influence of grain-size on the concentration/turbidity relation?
You are right. We have modified the text accordingly. In fact, the data at our disposal show that the distribution of sediment size is bimodal, with two peaks. The first one, around 5 micrometers, corresponds to the washload, and remains remarkably stable during the flood. The position of the second peak varies between 23 and 32 micrometers. Assuming that these data are representative, we expect no significant change of the turbidity-concentration relationship during a flood. We have modified the text to clarify this point. Given the data at our disposal, we cannot investigate in details the evolution of the grain size distribution during floods. In this context, the fact that direct manual measurements of the concentration of suspended sediment confirm the existence of hysteretic loops remains our best argument against the possibility of an artifact.

Line 324: I am not familiar with this chi2 measure of the quality of the fit. Could you clarify what it means when you say that chi2 ¡ 0.04 for the top 25% of your adjustments? Does it mean that the remaining 75% of the adjustments are not significantly different from a random model?
We acknowledge that this sentence is misleading. It is actually inherited from a preliminary investigation of the methods of inversion of our model. The bottom line is that the adjustment procedure works well, and that the model correctly fits the observation. So, for the sake of clarity, we removed this sentence and refer to the summary table of our results.

Line 535: I am surprised to read that the model provides a better estimation of the sediment flux than a rating curve for cases with clockwise hysteresis. How was this rating curve calibrated? Was it based on turbidity and discharge measurements for the event only, or over the full year? Could this rating curve be biased by the fact there are more measurement points in the falling than in the rising part of the hysteresis?

We constructed the rating curve from direct measurements of the concentration of suspended sediment in water samples collected by an automatic sampler all through the year. The automatic sampler is equipped with a pressure probe, which triggers the collection of 24 water samples, one every 15 to 60 minutes, as soon as the rivers stage exceeds a threshold set by the operator. In the dataset at our disposal, this threshold varies between 30 and 80 cm, that is way above the threshold of entrainment in the Capesterre river. Because of this, most measurements indeed belong to the falling rather than to the rising part of the hysteresis, as you suggest. This biais likely explains why the rating curve overestimates the concentration of suspended sediment in the river. We have modified the corresponding paragraph to clarify this point.