**Hydrology and
Earth System
Sciences
Discussions**

Interactive comment on "Parameter uncertainty and sensitivity analysis in sediment flux calculation" *by* B. Cheviron et al.

Anonymous Referee #1

Received and published: 22 March 2011

General comments

Against the background of in-stream sediment load estimation, the authors fit 4 concentration-discharge (C-Q) relationships to 21 long-term USGS datasets of daily resolution, aiming at optimal parameter sets using PEST. In doing so, the authors explore the effect of C sampling frequency by down-grading the daily datasets to 8 low-resolution frequencies, and also explore the effect of dataset length by truncating the original datasets down to 5 lengths. All scenarios are explored in combination. The authors also explore the effects of 20 combinations of positive and negative relative C and Q bias of 2 magnitudes. Each bias is applied equally to all points in a given dataset. Finally, the authors extrapolate their uncertainty assessment from the USGS data to a case study in France that was subject of another paper in press (Delmas et

al. 2011a).

While the topic is relevant, it has been investigated a lot, which makes it difficult to contribute much novel research. This paper, too, suffers from lack of novelty on top of some methodological choices that severely limit the practical relevance of the research. I cannot, therefore, recommend the manuscript for publication in HESS.

Specific comments

The introduction could be embedded in a wider soil erosion context and linked with some more recent work on C-Q relationships, similar to those models proposed here, and data uncertainties, e.g.: Krueger et al. (2009), Bilotta et al. (2010), Eder et al. (2010).

Some of the theoretical elaborations remain shallow and don't contribute much to the overall discussions. P1471, L9-12, L15-18, L26-28 – what is meant? P1471, L18 – the error term is implicit in any curve fitting procedure, it would be more interesting to discuss the realism of it! P1474, L4-P1476, L12 – this whole section (except P1475, L1-5) is unnecessary and makes things more complicated than they are. P1478, L1-3 – what is meant?

The materials and methods section is not clear as to what exactly was done. It would be helpful to review the models and their justification, especially model 4! Why the choice of PEST? P1475, L4 – do the 20 replicates mean the perturbed data scenarios? Why fixing parameters in models 3 and 4? The subsequent discussion about this remains dubious and seems to be blowing up a result that should have been avoided.

P1473, L14: "time-averaged" – mostly instantaneous, no? Or is the US data time-averaged?

One major disadvantage of the study is that the existence of C-Q relationships is never questioned. This is well known, see Bilotta et al. (2010) for a recent example. The median R^2 values in Table 2 of 0.58-0.65 might be classed as good or bad depending

on one's tradition, but how do the data look? How does the fit look? Do the residuals meet the implicit regression assumptions? Is there an optimal parameter set or would a joint parameter distribution better be used instead? By the way, the original Delmas et al. (2011a) paper doesn't answer these questions either!

Another major disadvantage is that the authors only consider data bias which they apply equally to all data points – the data are effectively just scaled along the C or Q axis or both. This leads to trivial results, e.g. P1478, L11-12; Figure 4b; P1479, L20-22; P1480, L24-26; Figure 6b. How realistic and therefore relevant in practice is a constant relative bias anyway? More prevalent is noise, which could be investigated by stochastic perturbations of the data points, including or excluding auto-correlation depending on the authors' justification. The main conclusions, then, were either "expected" (P1484, L23) or "known" (P1485, L1).

P1478, L22-26: Need to show results if this discussion is to feature. A "plausible hypothesis" doesn't seem satisfactory as a discussion though.

P1481, L16-17: Is this result shown? I must say I'm lost from here on, sorry, largely due to the convoluted wording.

P1485, L21-23: Despite the low expected error, how valuable is sediment load estimation over such long time?

Figure 3: Is this showing 1 scenario out of 21 datasets and 20 data perturbations? If so its generalising power is somewhat restricted.

Finally, I don't think one can extrapolate the uncertainty assessment from 21 USGS datasets to the French case study that easily (section 3.4). At best this can serve as an initial estimate – but how to update this for the new situation?

Technical comments

The style of the manuscript is somewhat convoluted which, together with some grammatical errors, makes it hard to follow. Some wrong wording: P1470, L5: resist to L8

and elsewhere: inferior to = smaller than L18: suffer from L22: ensue L25: pleading P1471, L2: inclines to L19: inaugural L27: somewhat efficient P1472, L16 and elsewhere: fitting LL17: into details = in detail P1474, L20: relevancy = relevance P1478, L16: distant and attenuated L21: silence P1480, L26: to the exception P1481, L10: traduce P1482, L11: is preferentially read in columns P1483, L4: acquainted

References

Bilotta GS, Krueger T, Brazier RE, Butler P, Freer J, Hawkins, JMB, Haygarth PM, Macleod CJA, Quinton J. 2010. Assessing catchment-scale erosion and yields of suspended solids from improved temperate grassland. Journal of Environmental Monitoring 12(3): 731-739.

Eder, A, Strauss, P, Krueger T, Quinton JN. 2010. Comparative calculation of suspended sediment loads with respect to hysteresis effects (in the Petzenkichen catchment, Austria). Journal of Hydrology 389 (1-2): 168-176.

Krueger T, Quinton JN, Freer J, Macleod CJA, Bilotta GS, Brazier RE, Butler P, Haygarth PM. 2009. Uncertainties in data and models to describe event dynamics of agricultural sediment and phosphorus transfer. Journal of Environmental Quality 38(3): 1137-1148.