We sincerely thank the three anonymous reviewers for their insight full comments on our manuscript. We admit that at one hand we took the issue of reproducibility a little too seriously, in the sense that our manuscript presents too many details, which much better fit into a supplement. We will revise the manuscript accordingly, and thereby shorten, streamline and focus the story line.

On the other hand it is evident that the new manuscript needs to better explain and motivate the key points why ***a) we add a new beast to the zoo****,* ***b) we think its parameters could and should be identified along a hierarchy of scales*** *and* ***c) we think that the proposed dataset is sufficient for model verification*** *(or to be precise to test the hypotheses inherently underlying questions a) and b). We apologize for being too brief in this respect and will better explain these key points in the revised manuscript as outlined below.*

a) Why adding a new beast and what shall it be good for?

In fact we did not invent a new beast but broadened the scope of an old beast named CATFLOW (born in 1997). CATFLOW allows numerical simulations of the water cycle in the critical zone from the plot to the catchment scales; for instance i) to explore the role of system structure and heterogeneity or the impact of uncertain initial states or rainfall (Zehe et al. 2001; Zehe et al. 2005; Zehe et al. 2013) and ii) to predict transport of dissolved substances ranging from tracers to pesticides (Wienhoefer and Zehe 2014; Klaus and Zehe 2011). By including erosion and sediment transport into CATFLOW we aim to provide a tool for a comprehensive assessment of nutrient and pesticide impact on surface water quality at the catchment scale, e.g. to explore flow path interactions and the role of facilitated pesticide transport and underlying controls in this context. The scope of CATFLOW-SED is thus clearly on robust process based predictions of sediment yields into surface water bodies from the headwater scale upwards, which are sensitive to land use, soil, topography and rainfall and reflect the threshold nature of erosion. This scope implies the challenge to balance necessary complexity, to avoid oversimplification, with parsimony to reduce model structural uncertainty. Key simplifications to achieve parsimony are i) our not accounting for rills and ii) the merging of the attacking forces of rainfall and overland flow by means of the vector sum of shear stress and rainfall momentum flux. We propose that these assumptions are well justified by the model scope and corroborated (in a sense that they cannot be rejected) by the good accordance of our simulations with observations of i) plot scale erosion and ii) sediment loads at the catchment outlet.

b) Why assuming a separated identifiability of the "erosion parameters" at a scale hierarchy?

Robustness implies a low model structural uncertainty and thereby equifinality, which is not at all restricted to conceptual models. Equifinality (Beven 2006) is i) inherent to governing equations, because of the interactions between gradient and resistance terms in these equations (Zehe et al. 2014) and ii) might arise from spatio-temporal interactions of different, partly independent processes controlling an integral system response at larger scales. For instance a doubled amount of detached particles within the upper parts of a hillslope will result in a similar sediment loss to the river, when being compensated by a twice as high sedimentation. As particle detachment and overland transport capacity depend both on flow velocity there is an obvious source for equifinality when determining the related parameters based on a single information/data source. Similarly, Klaus and Zehe (2010) showed that 13 different pairs of spatial density and hydraulic conductivity of macropores yielded similar inflow into a tile drain and thus equally acceptable model performance with respect to tile *drain discharge. Equifinality may, however, be reduced by fixing a single parameter in a parameter team, as shown for instance by Bardossy (2007) for the Nash cascade, or by fixing the hydraulic conductivity of a macropore (Klaus and Zehe 2010).*

Our study is based on the hypothesis/proposition that plot and hillslope scale sediment losses are essentially limited by different processes and underlying controls. Equifinality in the model may, thus, be reduced by estimating the parameters of different processes governing catchment scale sediment yields using essentially different data sources obtained at those scales where the corresponding process is limiting local sediment loss. More specifically we propose that i) the integral sediment loss from small plots is limited by particle detachment and underlying controls while ii) sediment loss from the hillslope to the stream is additionally limited by the transport capacity and sedimentation. We think that these assumptions are justified because sediment detachment and deposition are controlled by different forces (viscous momentum transfer and gravity) and time scale of sediment detachment is shorter than the time scale of deposition. This implies that depending of overland flow depth and velocity (and particle sink velocity), suspended particles get deposited after certain downslope transport distance and transport capacity is not limiting sediment loss at lower extents. In the revised manuscript we will highlight that the plot scale rainfall simulations were designed such, that integral soil loss was not affected by deposition within the plot (Scherer et al. 2012). The story line of the revised manuscript will put much more emphasis on the outline key hypothesis and on the suitability of our data set to test and reject this idea. Ultimately, we think that our simulation results do not allow rejection of our hypothesis i) because the model performed in very good accordance with observed plot scale surface runoff and sediment losses and ii) (in case of matching the observed hydrographs well) excellently with respect to the observed sediment loads at the catchment outlet.

As already stressed in the first manuscript the good model performance depends essentially on the good match of the observed hydrograph/runoff at both scales. At the catchment scale we fitted the model separately for the three rainfall runoff events. For the revised manuscript we will run the model continuously, by fitting it on the largest rainfall runoff event and work out the differences in simulated sediment loads compared to an individual fit. We will additionally provide the sensitivity of sediment loads with respect to changes in rainfall intensity. Behavioral soil parameter sets for rainfall runoff simulation are conditioned by the spatio-temporal resolution of rainfall data, as shown by Arnaud et al. (2002) and by Zehe et al. (2005) for CATFLOW. Hence, re-running the model with highly resolved rainfall input without re-adjustement soil parameters makes thus not much sense. In case they are re-adjusted to fit observed rainfall runoff, we expect the effect on sediment yields to be rather small.

c) On model verification and suitable data sets

We totally agree with Reviewer #1 that there are not data available which allow verification of environmental models. This is because we cannot verify models/proof their correctness, we can only falsify/ show that their underlying assumptions are (either in the sense of Popper (The Logic of Scientific Discovery 1959) or of Beven 2006) consistent with our experience/ data. CATFLOW-SED is as any model a technical implementation of our hypothesis, and we think our dataset is suitable for rejecting this hypothesis. This is because the model performs acceptable against integral sediment load data without adjustment of the detachment parameters. From this we can of course not conclude, that the model is "true" in a naive sense – we cannot reject it in a sense of Popper, or it is behavioral in the sense of Beven.

We will better stress this in the revised manuscript. Of course we would love to test the proposed process representations against more crisp data (for instance on sediment sources and sinks or on size sorted deposition within sediment source areas), because the model predicts these things. We are happy to do that if somebody provides such data.

Anonymous Reviewer #1

The paper represents a further effort to develop "simpler, but better" erosion models. While many reasons support such efforts, the paper does not represent a novel contribution, but just, to stay with analogy used by the authors, generates another "beast in the zoo of erosion models". There are three main reasons for this assessment.

1) The rationale for the algorithms eventually used in the model is not based on available empirical evidence to support the selections made. The lack of reference to data on e.g. C or P erosion or movement of pollutants illustrates this problem, i.e. model design choices are made, but not based on an argument derived from data published in the literature.

Please see our response to the major point a) 'Why adding a new beast and what shall it be good for?' on page 1 of this response.

2) The model is tested based on observations in the Weiherbach catchment that ignore the interdependence of scale and simulated process. Neither small-scale rainfall simulation, nor 69 m plots, nor sediment load at gauging stations reflect soil erosion, but sediment export from natural or artificial catchments with an increasing complexity of interaction of surface processes and properties. While the authors rightfully argue that models cannot and do not have to reflect such complexity, they do not provide evidence or a good rationale why discoupling scale and process interaction is suitable for using simple models. One would actually expect the opposite, i.e. that simple models work best if suited to a particular scale and geomorphic complex. This way, scale can be matched to a small number of processes dominating the erosional redponse observed on this scale and thus a simple numerical representation of erosion. In this manuscript, no argument matching scale to processes is made, which renders the model mostly an excercise of fitting observations to a set of equations perceived, but not shown to be relevent.

Please see our response to the major point b) 'Why assuming a separated identifiability of the "erosion parameters" at a scale hierarchy?' on page 1-2 of this response.

We would like to thank Reviewer #1 for pointing out, that our data reflects sediment export at various scales. This is in fact what we are aiming for. We apologize, that our arguments in the manuscript were misleading and will clarify this issue in the revised version of the paper as outlined in the major point a) on page 1 of the response.

Furthermore, we would like to point out, that <u>neither</u> of the parameters used in the approaches to quantify detachment, transport and deposition was fitted during the verification procedure. We have only calibrated the macroporosity factor f_m to match the observed surface runoff. This was done, because we wanted to exclude uncertainties caused by the simulation of runoff during the verification procedure of the approaches used to represent detachment, transport and deposition. Although neither of the parameters was fitted, the model reproduced the observed data on sediment export on all scales very well.

3) Finally, simple models with a strong empirical component should rely on a data base that represents the range of possible events. The data used in this paper appear to lack such quality of a good sample. For example, conducting rainfall simulation during spring and fall to collect data on different soil moisture conditions appears to ignore other differences, e.g. soil density and roughness which also change over time. The same is true for the assumptions made on sediment properties through the duration of movement through the catchment or effects of rainfall patterns and their impact on sediment loads, e.g. those identified by Quinton in "P-Erosion, does event size matter?". Overall, I do not think that the approach presented in this paper provides better modeling capacity than existing models, despite their shortcomings. Using the data from the Weiherbach catchment for a critical evaluation of the shortcomings of existing models would instead be much more desirable to prepare the way for the development of sound and capable simple erosion models.

Please see our response to the major point c) 'On model verification and suitable data sets' on page 2 of the response.

Furthermore, please note, that the rainfall simulation experiments did not only cover different soil moisture conditions but also different conditions of surface roughness and macroporosity as well as crop types. Details are provided in Scherer et al. (2012). We will clarify this issue in the revised version of the paper.

Anonymous Reviewer #2

This paper addresses the topic of distributed and physically based soil erosion modelling with a multi-scale approach that ranges from the plot scale to the small catchment scale (~ 3 km^2). The study focuses more specially on the issue of the parameterization of soil erosion defending the concept of the right balance of process parameterization to avoid over-simplification and model simplicity. My opinion is that the topic is very timely and the approach is very interesting. However, the conclusions reached by the paper are somewhat disappointing compared to the expectations that one can have on this subject. The main result is that the model reproduced well the observed data but the validation process seems incomplete.

Please see also our explanations on the major points a) - c) on Page 1-2 of the response.

I noted several points that need improvement. These main points are as follows:

1) The scope of the paper is too wide and the paper is a little too long, posing the risk of losing the reader. Many things are considered in the paper: the authors presents a new erosion model, they present a set of unpublished data, they make the calibration of the model, they try to evaluate the model and to make a sensitivity analysis and finally they use the model as a tool for prediction to test different land-use scenarios. It seems too much for a research paper and it affects the fluidity of the article and the coherence of the speech. This comment also applies to the title since the authors emphasize first the predicting capacity of the model related to land use changes and then they focus on the development and validation of the model across scales. I believe removing the part which deals with the prediction of different land use scenarios with the model would be relevant, especially since there is no data to verify these simulations.

We absolutely agree that the paper is too long and we will shorten and streamline it as outlined above. It is further true, that there is no data available to verify the two land use scenarios. However, the model was verified for different land use patterns, since the land use in 1994 and 1995 was different (Table 2). The model reproduced the sediment loads for the different events in 1994 and 1995 very well (Table 6). Please note that we have only calibrated the hydrographs (by adapting the macroporosity factor) for the events and <u>*no*</u> *erosion parameter was calibrated during the verification procedure of the model. Furthermore, we have only rearranged the given percentages of the land use categories in 1994 for the two scenarios (p. 3554, L. 1-5). We therefore conclude that the model is well validated and that the results for the two land use scenarios are thus realistic.*

2) I think it is more appropriate to limit the paper to the issue of the parameterization of soil erosion in the model based on the calibration-validation at various scales and to the issue of simplification in the selection of processes and their formulation. However, I feel that the part dedicated to the calibration of the model parameters is not completely treated. Some parameters such as the erosion parameters P1 and P2 have not been taken into account in the analysis and the justification of this choice is absent. On the other hand, one wonders why the authors have not used a framework such as the GLUE methodology to calibrate the parameters of infiltration and erosion processes. Nothing is said about the spatial distribution of rain in the methodology.

We have not taken the parameters p_1 and P_2 into account, because the derivation and parameterization of the detachment approach was presented in detail in Scherer et al. (2012): At first, a data set of laboratory experiments carried out under varying conditions of rainfall and overland flow (published by Schmidt 1996) but using resembling loess soil samples was used to determine the parameters p_1 and P_2. Secondly, data of the 58 rainfall simulation experiments performed on the 24 m^2 plots in the Weiherbach catchment were used to parameterize the erosion resistance f_{crit} under field conditions (please see also Section 3.2.3). In addition the parameter variability was tested in a split sampling approach. We will clarify this in the revised version of the paper and add details on the parameterization of equation 4.

Spatially distributed data on rainfall is available for the Weiherbach catchment (p. 3546, L. 6-7). However, in a first step we used the data of the meteorological station (WB0 in Figure 2) for the simulations at the catchment scale, which reproduced the hydrographs at the catchment outlet very well. Using the data of all available stations requires a regionalization procedure. Zehe et al. (2005) showed the influence of different regionalization approaches for rainfall and soil moisture data on runoff. Therefore the use of spatially distributed rainfall data requires a discussion on the uncertainty that is introduced by various regionalization methods. As this would go beyond the scope of the present manuscript, we are planning to use spatially distributed rainfall data in a future study.

We further refer to our general point b) at the beginning of the response, where we have explained the procedure of parameter identification at a hierarchy of scales.

3) The data used for this work are multi-scale and involve both experiments and observations, it is a very positive point. It is also interesting to promote the use of data collected 20 years ago to calibrate and evaluate the model. The multi-scale data involve rainfall experiments at the plot scale, observations of overland flow and erosion at the hillslope scale and observations of overland flow and erosion at the small catchment scale.

Nevertheless the extent of the data and the diversity of the selected events are open to criticism. For example, the data used at the hillslope scale are cumulative exports of eroded particles at the event scale and are limited to a single event. This seems insufficient to confidently assess a distributed model of this type.

Of course, it would be much better to have a data set that represents a variety of events. However, during the monitoring period of the long term soil erosion plot, only one storm event was sampled. Nevertheless the cumulative erosion volume of this event was reproduced very well by CATFLOW-SED. Please note, that for the simulation of the storm event on the long term soil erosion plot, no calibration step was performed: nor for the simulation of runoff, neither for the simulation of the erosion volume.

4) Finally, I do not understand why the authors consider the multi-class approach in the description of the model and do not say anything about it in the rest of the paper, nor in the presentation of the experimental dataset or in the calibration and evaluation of the model. It seems contradictory with the idea to try to simplify the model. Why do they take into account several soil classes without mentioning specific parameters to each class and without evaluating the model results with field data? Is there an interest to preserve the multi-class approach in this study?

Thank you very much for pointing this out. We used a multiclass approach because we are planning to implement also the particulate transport of contaminants within CATFLOW-SED. Therefore the consideration of various particle size classes is indispensable. The main parameters for the multiclass approach are the mean percentage of grain size fractions (provided in Table 4) and the mean particle diameters of each class which are defined by the classification scheme of soil texture (will be added in Table 4 in the revised manuscript). We will clarify this in the revised version of the manuscript. Furthermore, we will provide details on the representation of the multiclass approach in section 2.2.2 'Coupling of water and sediment budgets in CATFLOW-SED' as well as in the discussion section.

5) A list of symbols is missing in the paper to help the reader navigate.

We will include a list of symbols in the revised version of the paper

The following comments regard more specific points:

- p.3531, L.6-8 : this sentence is not clear. Please re-formulate

This hypothesis is related to the scale dependent verification procedure (please see response to major point b). We will clarify this in the revised version of the manuscript.

- section 2.1.1 The authors refer to detachment of soil particles and aggregates but they forget to specify that it is detachment by flow. For instance, p.3533, L.14 or p.3535, L.10.

Will be added. In addition we will shorten the detailed discussion on the interaction model of the sub-processes of erosion in the revised version of the manuscript.

- Eq(4) p.3536 : I do not understand why the detachment by rainfall impact and the detachment by flow are dealt in a bi-linear way. This expression means, for example, that if the energy of rainfall is sufficient to overcome the resisting forces of the soil, the soil can therefore be eroded by overland flow even if the shear stress of the flow is very low and inferior to the resistance of the soil. It is surprising as detachment by flow is usually conceptualised as a threshold process in cohesive soils.

We used a bi-linear approach, since overland flow in the underlying water balance model CATFLOW is represented by sheet flow and no distinction is made between interrill and rill areas (see section 2.2.1). In case of rainfall, more particles are detached by flow than without rainfall. This is proven by rainfall simulation experiments in the laboratory (please see Scherer et al. 2012 for details). In theory it is possible, that in case of a heavy storm event, the resistance of the soil against detachment (erosion resistance f_{crit}) can be overcome by the attacking forces of rainfall alone. However, such a heavy storm event in a loess catchment is usually associated with the occurrence of overland flow, since Hortonian overland flow is the dominating process in such areas. Theoretically it might be possible, that the rainfall intensity could be high enough to overcome the erosion resistance of the soil f_{crit} without producing surface runoff. In this (rather theoretical) case, the detachment rate of particles is quantified by the model, but without overland flow, no transport capacity is available to translocate particles further downslope and the resulting erosion rate will be zero.

- Eq(5) p. 3539: it is not clear how Tc is calculated

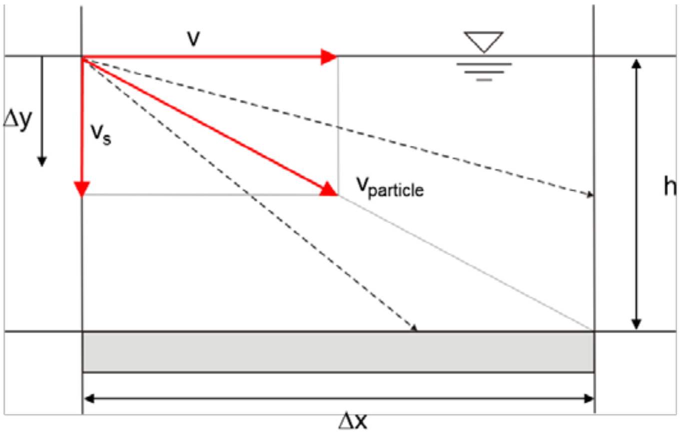
We displayed the equation to quantify sediment transport capacity using the dimensionless variables "transport intensity" ϕ and "stream intensity" θ. Dimensionless variables are commonly used in hydraulic engineering. To quantify the transport capacity, one has to insert the equations for all variables, which are ϕ, θ and λ and then solve the equation towards the transport capacity T_C. We will provide the resulting equation in the revised manuscript.

- p. 3539, L. 12-15: what is the relevancy of presenting a multi-class approach if it is not discussed in the rest of the paper?

This section refers to the scale dependent verification procedure that was used to test the representation and parameterization of the various sub-processes of erosion. Please see our answer to the general point b) at the beginning of this response. We will clarify this issue in the revised version of the manuscript.

- p. 3541: the discussion about the Cdep coefficient is not clear. What is this coefficient for?

The sinking velocity of each particle class is quantified using equations 6-9. The following Figure illustrates the relation between flow depth h, flow velocity v and sinking velocity v_s in a model discretization element of the length Δx. Δy is the vertical travel distance of a particle.



Δy is quantified by using Eq. 10. Unfortunately we realized that equation 10 contains a typing error: v_p should be flow velocity v:

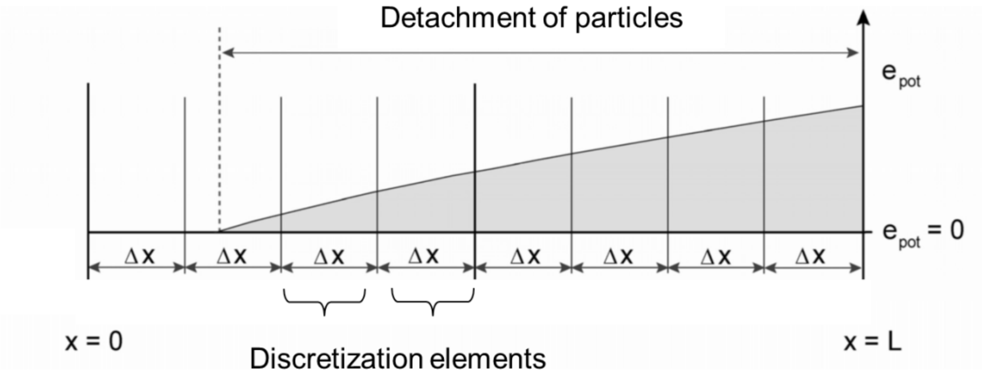
$$\Delta y = \frac{v_s}{v_p} \cdot \Delta x \rightarrow \Delta y = \frac{v_s}{v} \cdot \Delta x \quad \textit{(will be corrected)}$$

A particle of a certain particle class can only be deposited within the discretization element Δx, when it reaches the ground in a certain time step Δt. In the case that $\Delta y \geq h$ for a specific particle class, all particles of this class will be deposited. In case that $\Delta y < h$, only a share of this particle class will be deposited. This share is equivalent to the deposition coefficient C_{dep}, and is quantified by the relation of Δy / h (Eq. 11).

We will clarify the meaning of C_{dep} in the revised version of the paper according to the given explanations above. In addition we will make clear, that deposition is quantified for the various particle classes, which is missing in the current version of the paper. This will be also added in section 2.2.2 'Coupling of water and sediment budgets in CATFLOW-SED'.

- section 2.2.2.: Eq(15): what is the purpose of this integration? What do you mean by discretization element exactly? It is surprising that you do not talk about multi-class modelling in this section.

The following figure displays the spatial development of the potential detachment rate e_{pot} along a slope of length L. A Discretization element is a model element of length Δx.



The potential detachment rate e_{pot} (kg m^{-2} s^{-1}) is quantified by equation 4:

$$e_{pot} = p_1 \cdot (\tau + P_2 \cdot m_r - f_{crit}) \qquad if \;\; e_{pot} < 0, e_{pot} = 0$$

Where m_r is rainfall momentum flux (N m^{-2}), τ is shear stress (N m^{-2}), f_{crit} is the erosion resistance (N m^{-2}) and p_1 (s m^{-1}) and P_2 (-) are empirical parameters.

Due to the non-linear character of shear stress τ, the potential detachment rate e_{pot} is non-linear, too. In addition, e_{pot} is zero, in case that the erosion resistance f_{crit} is not overcome by the sum of the attacking forces of flow (quantified by shear stress) and rainfall (quantified by the momentum flux). Because of the non-linear character as well as the threshold character of e_{pot}, the **total** *detachment rate per unit width $q_{s,pot}$ (kg m^{-1} s^{1}) for each discretization element has to be calculated by integrating the area-specific potential detachment rate e_{pot} along x (equation 15).*

$$q_{s,pot} = \int_{x_{i-1}}^{x_i} e_{pot} \cdot dx$$

We will clarify this in the revised version of the paper and we will also describe the approach of multiclass modelling in section 2.2.2.

- p. 3547, L.7: burrows instead of borrows

Will be corrected

- p. 3547, L. 22-23: the date and the rainfall amount do not match with the data in Table 3

Sorry, for this mistake. The rainfall amount in the table is correct. We will adapt the rainfall amount in the text.

- section 3.3 : that would be much easier if you would be more precise about which data are used for calibration and which data are used for validation. You could make a Table that explains the distinction between calibration and validation and the method used in each subsection of sections 3.3 and 4.

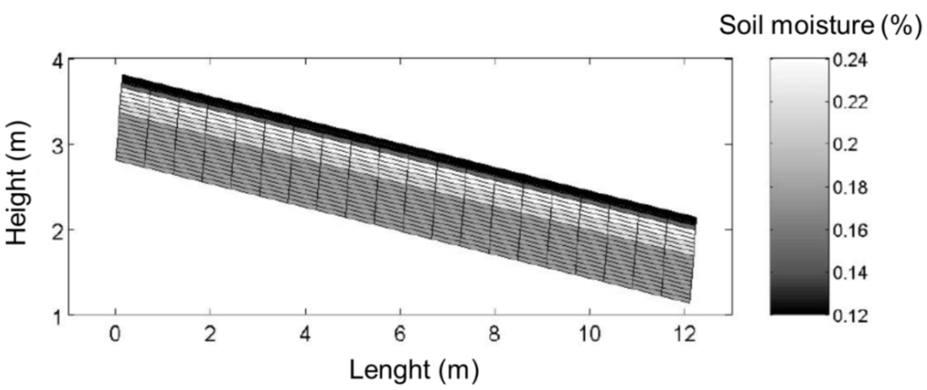
Thanks for pointing this out. It is a very good idea.

- p. 3550, L. 7-10: these are hypothesis but they are not proved

This section is related to the scale dependent verification procedure that was used to validate the sub-processes of erosion. Please see our response to the major point b) at the beginning. We will clarify this in the introduction of the revised manuscript.

- p. 3551, L. 2-4: the discretization is not clear to me, in the 3 dimensions.

It means the number of vertical (29) and lateral (21) discretization elements in the cross section of the plots for modelling the rainfall simulation experiments (as displayed in the Figure below). We will clarify this in the revised version of the paper.



- p. 3551, L. 10: why not to calibrate the Manning coefficient as well

The Manning coefficient of the irrigation plots was determined by fitting the falling limb of the observed hydrograph after irrigation was terminated (please see p. 3548, L 23-25). Since we had measured data for this parameter we used the data for modelling the experiments.

- section 3.3.3: there is no hydrograph and sedigraph at the plot scale to evaluate the model?

No, unfortunately not. On the long term soil erosion plot only the cumulative sediment volume was sampled.

- section 3.3.5: why do you perform a sensitivity analysis at the end? That should be done at the beginning and taking into account more parameters. Furthermore, the methodology used to carry out this sensitivity analysis is not clear.

We performed these simulations to analyze the uncertainty of the model parameters erosion resistance f_{crit} and macroporosity factor f_m on the plot scale. However, we think, that this part of the study is dispensable and will therefore show the variation of the model output by a continuous simulation as outlined in the major point b) on page 1 and 2 of the response.

- p. 3554, L. 1-18: I would remove this paragraph.

Please see response to the second major point of Reviewer #2

- p. 3555, L. 11-13: hopefully the agreement is good as it is the calibration

Only the macroporosity factor for the simulation of runoff was calibrated. For erosion modelling, measured parameters were used and no further calibration for erosion was carried out. We will clarify this in the text.

- p. 3555, L. 17-29: I do not understand the meaning of these regressions? Why do you perform this statistical analysis? Is that for the regionalisation of the parameters used in section 4.1.3? This is not clear to me.

We will clarify this in the revised version of the paper.

- section 4.1.2: only one event with cumulative data at the event scale. It seems not sufficient to evaluate such a distributed model.

We absolutely agree. We would be very happy to have additional data at this scale. Unfortunately, we do not have it.

- section 4.2.2: I would remove this section

Please see response to the second major point of Reviewer #2

- section 5: Be careful with the terminology. It is important not to mix calibration and validation. Moreover, it is preferable to say “CATFLOW-SED reproduced the sediment loads” instead of “CATFLOW-SED predicted the sediment loads”

We agree. Thank you very much for this advice.

- p. 3561, L.22: burrows instead of borrows

Will be corrected

- Figure 1: term "precipitation (qlat)” shouldn’t be in the middle of the figure?

Figure 1 will be adapted

- Figure 2: that would be fine to add the exact coordinates of the site

Coordinates will be added in Figure 2

- Figure 2 and 4: could you add a soil map or a map of soil parameters (soil resistance for example) to one of these figures?

The pattern of the soil types in the Weiherbach catchment is very regular as shown in Figure 3. We will check if it might be better to display a soil map in Figure 4 instead of Figure 3. Thanks for pointing this out.

Anonymous Reviewer #3

This paper focuses on the topic of spatially distributed and physically based modelling of runoff and erosion on multiple scales. This is an interesting and important topic, especially since many models exist which all have their advantages and disadvantages. This paper promises to unfold a new model that is simple, but physically based, with measurable parameters and also applicable over multiple scales. This raises high expectation which are, unfortunately, not fully met in my opinion.

Most importantly, it is not completely clear what this particular new model adds to the already existing 'zoo of models'; is it not another 'animal in that zoo' to speak with the terms the authors use? Although the model seems to do well, a more in depth discussion on this issue is needed, clearly pointing out the advantages of this model as compared to existing ones. Related to this is the question: did none of the existing model simulate the hydrological and sediment redistribution processes satisfactorily for the catchment?

Please see our response to major point a) 'Why adding a new beast and what shall it be good for?' on Page 1 of this response.

Secondly, the model is referred to as 'simple' or, at least, as simple as possible. It remains subjective what simple is, but in my opinion, the model is far from simple, including many processes and equations. See for an interesting discussion on simplicity of models e.g. Paola (2011). Related to this is the number of parameters this model has. The aim of the authors to include only really 'measurable' parameters (as opposed to lumped ones) is indeed something to aim for. However, this model needs very many parameters. In the catchment that is being simulated, a very detailed, extensive and long measurement campaign has been carried out, which provides this data. However, for most catchments and studies, such amount of data is not available and it is often too expensive to obtain. So, how feasible or useful is this model for general use in other, more data-poor, regions?

We apologize that our argumentation was misleading in this point. We did not develop a 'simple model', but a model that balances 'necessary complexity with greatest possible simplicity'. We will clarify this issue in the revised version of the manuscript (please see also the response to major point a) 'Why adding a new beast and what shall it be good for?' on Page 1 of this response. We have further developed a well-tested process based hydrological model to account for erosion and sediment redistribution, too. We agree that it is not possible to apply this model in data poor regions. However, it should be possible to apply the model in similar intensively cultivated loess regions and transfer the process parameterizations to such catchments. Nevertheless, data series on rainfall and climate runoff as well as information on land use patterns and soil textures are indispensable to run the model in similar catchments.

Thirdly, in the calibration and validation of the model, the biggest (three) storms are used. I agree that usually most erosion is caused by large storms, but still also quite some sediment redistribution might occur in smaller storms that are much more frequent. Why were smaller storms not tested?

The three storms that were used for the verification of the model at the catchment scale are indeed the three biggest events, but they were also the only ones which produced significant erosion rates during the monitoring period in the Weiherbach catchment. Thanks for pointing this out, we will clarify this issue in the revised version of the paper. The Kraichgau region, where the Weiherbach catchment is located, is highly susceptible to thunder storms in spring and summer. However, these thunderstorms occur very locally and only a few catchments in the region are usually affected. Although the affected areas are small, the erosion rates might be very high, as shown for the three events in the Weiherbach. This is the reason why only a few events were sampled in the Weiherbach, although the catchment was monitored over a period of several years.

Nevertheless, small erosion events have been observed in the Weiherbach. But for small events the processes causing erosion are highly variable and the basic model assumptions, such as the production of extensive surface runoff on the hillslopes, are no longer valid. For example, it may happen during a rainfall event that the water collects on the sealed rural roads and flows into an adjacent field causing a small erosion rill. Such a process can produce erosion, in case it happens close to the stream. However, it is not possible to model such a process with CATFLOW (or with any other model aiming at the catchment scale). The difficulty of modelling small erosion events was also pointed out by Jetten et al. (2003) and Nearing (2006), because of the random and local occurrence of various erosion phenomena. We will discuss this issue in the revised version of the manuscript.

A final major issue of the paper is its length; in my opinion the paper is too long and contains too many topics: a new model, its calibration/validation, sensitivity analysis, data collection and requirements and scenarios and effects of land use change. I think the latter is too much and should be skipped from the paper; or, alternatively, two consecutive papers can be prepared covering all these subjects. However, I also would like to compliment the authors; for such a long paper with so many subjects, it was still good to follow most of the time.

Some minor issues and questions are:

- Abstract: last sentence difficult to follow '..to mitigate...loess landscapes' (p.3529), please rewrite.

Will be clarified in the revised version of the manuscript

- 'emissions' (p. 3529 L7): rather say 'inflow' or comparable.

Will be corrected

- 'burrows' instead of 'borrows' (p. 3529 L21; happens more often, please check)

Will be corrected

- 'landscape' (p.3530 L19)

Will be corrected

- if this complies with the journal's regulations, please give references in chronological order; this is not always done.

Thanks for pointing this out. We will sort the references within the text in chronological order. This is compliant with the journal regulations.

- (section 2.2.1) The model uses 'model elements' to discretize hillslopes and the catchment; so it is not grid or cell based. What's the advantage of this? Is this also applicable to more complex catchments, as each hillslope section is assumed to act similarly? Many cellbased models are sensitive to resolution; how is this for this model?

This is due to the underlying hydrological model CATFLOW which uses discretization elements along hillslopes as displayed in Figure 6. This discretization scheme is applicable to more complex terrain, since the size of the hillslopes as well as the size of the discretization elements can be chosen individually. There is a "wizard" available for the spatial discretization in a GIS environment. Wienhöfer & Zehe (2014) have already applied CATFLOW in more complex terrain such as the Austrian Alps.

We have performed a sensitivity analysis for 9 hillslopes (length: 100 m, slope 10 %) with varying lengths of the discretization elements Δx. The first slope was discretized with Δx = 50m. For the other slopes Δx was halved until a length 19.8cm was reached. All 9 hillslopes were simulated using a 1 hour long rainfall with an intensity of 40 mm/h. The erosion resistance f_{crit} was chosen as 1 N m^{-2}.
Please find the results of the cumulative discharge (m³) and eroded soil material (kg) in the table below. The discretization error is quite low, ranging between 2-3% (for Δx=20cm to 50m) for both, the cumulative discharge and the cumulative eroded soil material.

Δx [m]	**50**	**25**	**12.5**	**6.25**	**3.125**	**1.56**	**0.781**	**0.391**	**0.198**
Cumulative discharge [m³]	18.53	18.67	18.76	18.82	18.83	18.86	18.98	19.05	19.06
Cumulative eroded soil material [kg]	7126	7200	7242	7265	7275	7280	7295	7303	7315

The very minor influence of the discretization length on discharge is due to the numerical representation of the diffusion wave equation by using an 'upstream' procedure. Regarding the simulation of erosion, only the process of detachment is sensitive to the discretization length Δx (please see also response to Reviewer #2, comment on section 2.2.2.: Eq 15). Therefore the discretization elements are further discretized into 'sub-elements' during computation as explained on p. 3544, L. 8-14. We found that a length of 1m for the sub-elements is sufficient to avoid discretization errors. We will clarify this in section 2.2.2 in the revised version of the manuscript.

- 'up to' instead of 'to a massive' (p. 3545 L7) - 'thick' instead of 'huge' (p.3545, L18)

Will be corrected

- consider showing the land use types in cake-diagrams instead of in a table (Table 2)

Will be corrected

- p 3556 L27: 'This value is slightly larger': which value? The values mentioned are actually the same?

We have referred to the result of the macroporosity factors of the plot scale experiments. We will clarify this in the revised version of the manuscript.

Ref: Paola C. 2011. In modelling, simplicity isn't simple. Nature 469: 38.

Thanks for the note on this reference. It fits very well into the context of the paper!

Additional References:

Arnaud, P., Bouvier, C., Cisneros, L., and Dominguez, R.: Influence of rainfall spatial variability on flood prediction, Journal Of Hydrology, 260, 216-230, 2002.

Bardossy, A.: Calibration of hydrological model parameters for ungauged catchments, Hydrology And Earth System Sciences, 11, 703-710, 2007.

Beven, K.: A manifesto for the equifinality thesis, Journal of Hydrology, 320, 18-36, 2006.

Jetten, V., Govers, G., and Hessel, R.: Erosion models: Quality of spatial predictions, Hydrological Processes, 17, 887-900, 2003.

Klaus, J., and Zehe, E.: A novel explicit approach to model bromide and pesticide transport in connected soil structures, Hydrology And Earth System Sciences, 15, 2127-2144, 10.5194/hess-15-2127-2011, 2011.

Klaus, J., and Zehe, E.: Modelling rapid flow response of a tile drained field site using a 2d-physically based model: Assessment of "equifinal" model setups, Hydrological Processes, 24, 1595 – 1609, DOI: 10.1002/hyp.7687., 2010.

Nearing, M. A.: Can soil erosion be predicted? , in: Soil erosion and sediment redistribution in river catchments edited by: Owens, P. N., and Collins, A. J., CABI Publishing (Wallingford), 145-152, 2006.

Scherer, U., Zehe, E., Traebing, K. and Gerlinger, K.: Prediction of soil detachment in agricultural loess catchments: Model development and parameterisation. Catena 90, 63–75, 2012.

Schmidt, J., 1996. Entwicklung und Anwendung eines physikalisch begründeten Simulationsmodells für die Erosion geneigter landwirtschaftlicher Nutzflächen (Development and Application of a Physically Based Simulation Model for Erosion on Slopy Cultivated Areas): Habilitation dissertation. Berliner Geographische Abhandlungen, 61. University of Berlin.

Wienhofer, J., and Zehe, E.: Predicting subsurface stormflow response of a forested hillslope - the role of connected flow paths, Hydrology And Earth System Sciences, 18, 121-138, 10.5194/hess-18-121-2014, 2014.

Zehe, E., Becker, R., Bardossy, A., and Plate, E.: Uncertainty of simulated catchment runoff response in the presence of threshold processes: Role of initial soil moisture and precipitation, Journal of Hydrology, 315, 183-202, 2005.

Zehe, E., Ehret, U., Blume, T., Kleidon, A., Scherer, U., and Westhoff, M.: A thermodynamic approach to link self-organization, preferential flow and rainfall-runoff behaviour, Hydrology And Earth System Sciences, 17, 4297-4322, 10.5194/hess-17-4297-2013, 2013.

Zehe, E., et al. (2014), HESS Opinions: From response units to functional units: a thermodynamic reinterpretation of the HRU concept to link spatial organization and functioning of intermediate scale catchments, Hydrology and Earth System Sciences, 18(11), 4635-4655.

Zehe, E., Maurer, T., Ihringer, J., and Plate, E.: Modelling water flow and mass transport in a loess catchment, Phys. Chem. Earth (B), 26, 487-507, 2001.