Dear editor, dear referee,

Thank you for this second opportunity to revise our manuscript. Please find below Anonymous Referee #2's comments (in normal font), followed by our responses (in bold) and the location of the changes we made in the manuscript (underlined).

Kind regards,
On behalf of all co-authors
Laurent Gourdol

Anonymous Referee #2

I would like to thank the authors for this revised version of the manuscript, and more specifically for the detailed answers to the initial comments. The authors have addressed a series of concerns on the originality of the study, and I must admit that their justification as to why this work is worth publishing in HESS is totally reasonable. I have no major objections anymore for this study to be considered for publication in HESS (after clarifying some remaining points of concerns). However, is worth mentioning that even though I'm satisfied with the author's rebuttal, the revised version of the manuscript remains in essence rather similar to the first version, and was not substantially modified, as one would have expected after a "major revision".

We would like first to thank Anonymous Referee #2's (hereafter AR2) for the additional time spent on reviewing our work and for providing thoughtful suggestions of improvement.

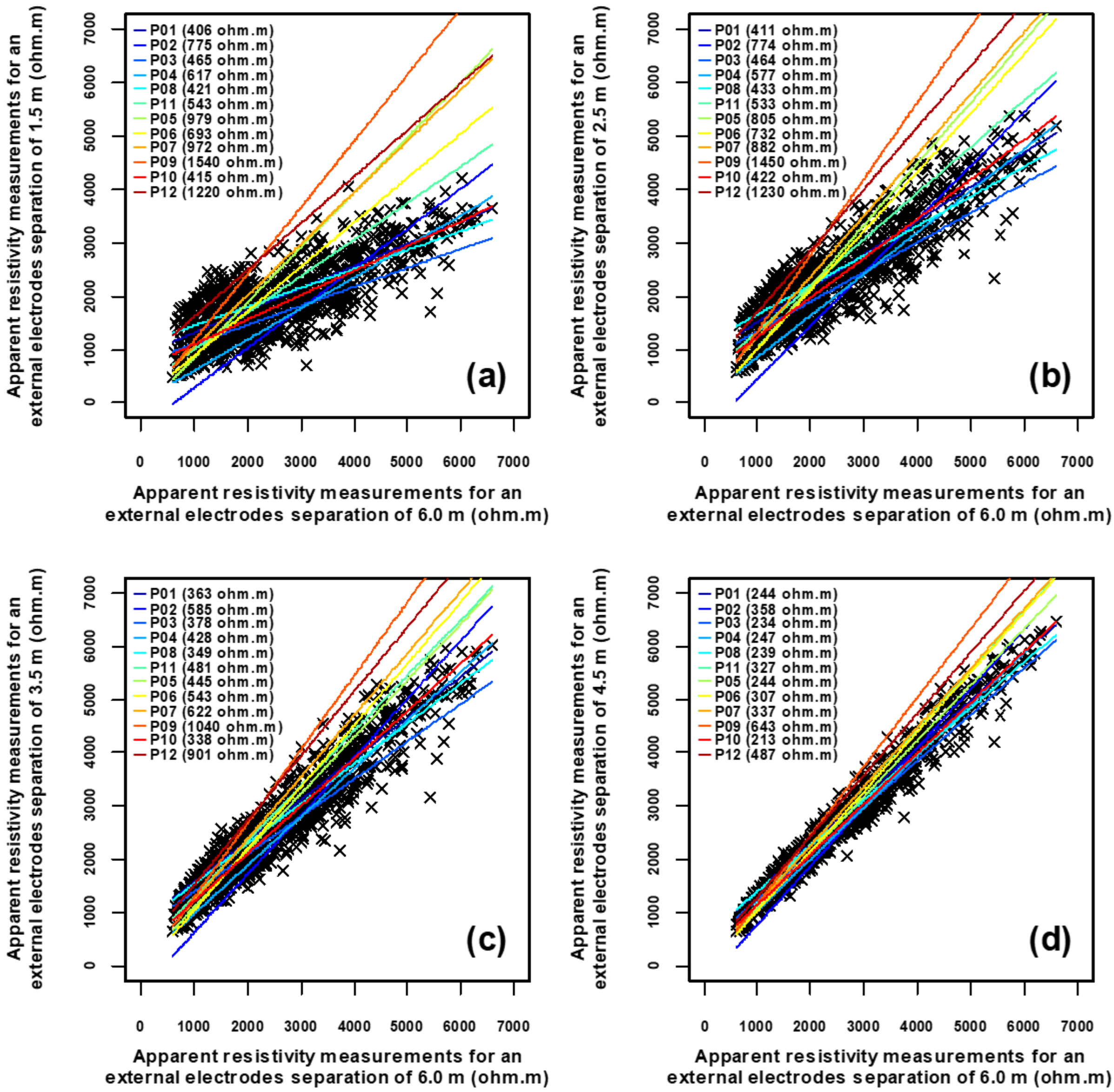
The proposed novel procedure for ERT measurements, which is the main focus of the manuscript, and described in Sect. 2.3 remains too poorly explained in this revised version of the manuscript, and I'm afraid is still subject to misunderstanding. The new figure 4 illustrating the novel approach is clearly a really good addition to help capturing the main concept, but, apologies for this, I still struggle to understand the method. The authors state in their answer to the editor that in the original version of the manuscript, the reader might miss a key step of the proposed procedure. The new version of the manuscript does not seem to bring much more explanation on this key step, as no improvement of the manuscript has been included in Sect. 2.3 of the methodology to support the new Figure 4. The missing key point concerns what data are used in the linear regressions. It is indeed not clear what data are plotted against each other on the scatter plots (Fig. 6 and Fig. 10). I know realise that this might come from a missing information in the first step detailed in Sect. 2.3, which states that "from the set of apparent resistivity data measured with an ES of 2 m, we extract the first acquisition level of apparent resistivity data (for the smallest possible external electrodes separation, i.e. 6 m). For this acquisition level, we extract – from the set of apparent resistivity data relying on an ES of 0.5 m – four subsets of apparent resistivity data for smaller external electrodes separations of 1.5, 2.5, 3.5 and 4.5 m, respectively." There is no information as to what these subsets are made of. Is it the apparent resistivity of the measurement array which has the pseudo-x position closest to that of the array with the larger ES? Is it an average of the apparent resistivities for all the arrays of a specific smaller ES comprised within the larger ES? As things stand, the reader can only guess what this subset is made of, and I would suggest the authors to explain in much more detail how it is extracted. It wouldn't arm anyway to have a slightly more detailed Sect 2.3 of the manuscript. This is the claimed main novelty of the study, so I would really strengthen the explanation of the method. In the current version of the manuscript, there is a rather shortly explained method, with lengthy examples, but if the method is not clear in the first place, the examples are not really helpful. Clarifying how the subsets are extracted in the legends of Fig. 6 and Fig 10 would also be welcomed.

We agree with AR2's assessment and admit that the explanation of our upgrading methodology in section 2.3 is not clear/detailed enough and might be consequently misleading for the reader. <u>According to AR2's suggestions, we revised section 2.3 as well as captions of Figure 6 and 10. We also modified the caption of Figure S14 in the supplement.</u>

A practical aspect of the proposed method could also be included in the discussions. Indeed, have the authors investigated the appropriate number of ERT profiles with smaller ES needed to improve significantly the accuracy of the ERT measurements with larger ES? It could be interesting to assess if using a smaller number of ERT profiles than the 12 field datasets gives similar results. Since the area seems to have a rather constant solum thickness, could only 1 ERT profile with smaller ES provides a robust linear regression? This aspect might be critical to assess the practical feasibility of the method, and is worth including in the discussion.

We thank AR2 for this good point. Unfortunately, we did not investigate precisely the impact of the number of ERT profiles using an ES of 0.5 m on the accuracy of the upgraded ERT profiles relying on the ES of 2 m for the Weierbach catchment. The 12 plot scale ERT profiles were selected based on our *a priori* knowledge, in order to have a set of spatially well distributed profiles over the catchment area, and covering the range of prevailing local geomorphological characteristics. However, as asked specifically by AR2, we can assert that using a single ERT profile with an ES of 0.5 m would be highly uncertain in the Weierbach catchment, despite the fact that this study area exhibits a rather homogeneous solum in terms of thickness and resistivity. The ability of the calibrated linear regressions relying on one single profile to deliver accurate prediction for the full data set is indeed highly variable from one profile to another as shown in the figure on the next page. Although some profiles lead to predictions comparable to those obtained from the complete dataset, others provide poor results. We nevertheless agree with AR2 that the number of accurate ERT profiles with the small ES (together with their location) is of key relevance for the upgrading procedure to succeed. On the one hand, as highlighted by AR2, if this one is too large, the applicability of the upgrading procedure might be cumbersome. On the other hand (and most importantly in our opinion), if it is too small, the upgrading of the ERT profiles relying on the large ES might lead to less reliable and inaccurate inverted results, or even biased results. Practically, for both cases, it is about evaluating the efficiency of the dataset used for calibration to deliver robust estimates of the surficial levels of apparent resistivity regarding the dataset size. <u>New elements have been added to the penultimate paragraph of the discussion subsection 4.2 in order to deal with this aspect, as well as a new citation has been added to the reference list.</u>

<u>Finally, it is worth noting that we indicated in our revised manuscript the contribution and competing interests of the authors after the conclusion. A distinct data availability section has also been created (the data availability was described in the Acknowledgments of the previous version of our manuscript).</u>



Scatter plots relating the apparent resistivity data corresponding to the first pseudo-depth acquisition level for an ES of 2 m (external electrodes spacing of 6 m) versus the shallower first four surficial apparent resistivity levels for an ES of 0.5 m with external electrodes separations of 1.5 (a), 2.5 (b), 3.5 (c) and 4.5 m (d) for the 12 plot scale ERT profiles measured in the Weierbach catchment. Each of the linear regressions presented in this figure (colored lines) was calibrated using one single ERT profile (one colour represents one specific profile). Their accuracy indicated in brackets (root mean square error) was however computed by considering all the data of the 12 ERT profiles.