First, we would like to thank Anonymous Referee #2 (hereafter AR2) for taking the time to read and assess our manuscript. Here below, we have listed all AR2's comments in normal font, followed by **our responses in bold**.

General comment

This manuscript presents an original study on the impact of electrode spacing (ES) on the resolution of resistivity models resulting from Electrical Resistivity Tomography (ERT) surveys. The study is illustrated by a series of synthetic data through forward modelling and by one field case study at the Weierbach catchment in Luxembourg. I must say, the technical aspects of this paper are excellent. The authors have used a state-of-the-art methodology and most of the processing steps of the ERT data, both for the forward modelling and the inversion, are relevant and well explained (some-times too much). In fact, this is a very good technical paper. However, the research question is trivial and the proposed updated methodology is somewhat questionable. The authors even mention in the abstract that (most of) the findings are obvious (!). Decreasing the ES will indeed result in a greater resolution of the resistivity model, no doubts about that. Now, this paper has the merit of documenting very well the effect of different ES via a synthetic case study (Fig 3 is a great one for teaching purposes for example), and to illustrate that decreasing the ES has also an effect on the precision of the retrieved boundaries at depth (which is potentially the most interesting outcome of the paper). But then the next question is: is this paper suited for HESS? I am not too sure, since it proposes a slightly questionable updated methodology for ERT measurements done on sites with very specific characteristics in terms of homogeneously flat soil/geological structures, and for those interested to image both the thin soil layer and deeper structures. I don't see that this falls into the scope of HESS to be honest. Sure the authors have oriented the introduction on the benefits of ERT for hydrological investigations, but the rest of the paper does not really matches with HESS at this stage, even when discussing the right way of estimating precise depths of boundaries between deeper layers. I think that the methodology would be more suited for publication in a more technical journal on applied geophysics, potentially focusing primarily on the synthetic modelling. And then perhaps, building up on the published methodology, the authors could demonstrate its benefit for hydrological purposes through a real case study. But again, the technical aspects of this paper are outstanding.

AR2 qualifies our work as technically outstanding and very well documented. We would like to thank her/him for that assessment. However, we also identify in AR2's general comment some concerns about our study that we want to address below.

AR2 states that our manuscript might be a good technical paper dedicated to geophysicists, but that it eventually does not perfectly fit to HESS. It is true that the technical aspect dealing with Electrical Resistivity Tomography (ERT) limitation/improvement is central to our study, but in our opinion the guiding idea behind our work – which is to provide new insights into subsurface mechanisms and aid in the parameterization of subsurface flow and transport models – is inextricable and just as important. For that reason, we believe that our research work must be considered as a hydrogeophysical study rather than dealing with a technical geophysical problem alone. As documented in the introduction of our manuscript, several authors have recently pointed out the subsurface as being the greatest knowledge gap in the understanding of hydrological processes, with a greater investment into "seeing" the subsurface needed to provide the Earth System Modelling community with critical guidance on how to parametrize model subsurface structure depths and properties. This general framework has strengthened our belief that the topic of our work is well-suited to HESS. We are thus convinced that our research work provides results and findings that are of interest to and can be used by the wide research community targeted by HESS (see Discussion section 4.3 From the Weierbach catchment perspective and beyond).

AR2 also suggests that our work might be somewhat trivial. Indeed, he/she states that the impact of ES on inverted ERT images accuracy is already a well-known issue. While we fully agree of course with AR2 on this statement, we nonetheless thought that it would be worth mentioning this obvious fact, while targeting a non-geophysicist readership that might be less familiar with this technique. But our work does

not focus on this general issue, but rather addresses a specific and less trivial aspect. As indicated in our introduction, we were wondering whether deep structures are well defined if the shallow structure is not well sampled. Our work was specifically concentrating on the influence of the ES on ERT results accuracy when a top thin layer is present in a subsurface structure (e.g., the soil layer within a typical soil-to-substratum continuum). To the best of our knowledge, no study has documented so far in detail from which ES threshold (as well as why and how) the accuracy of inverted ERT images is significantly affected for a typical regolith structure. Note that such a layered sequence of soil–saprock/saprolite–bedrock mirrors the subsurface of many natural contexts. We found out from our study that the thickness of the most superficial layer (i.e., the soil in our case) must be considered when choosing the ES – even when solely aiming for the characterization of deeper layers. We are convinced that this key result is of importance for the hydrological sciences community in search of accurate characterizations of regolith geometry and properties. Our approach ultimately untaps the potentially biased results (in terms of both resistivity distribution and interface delineation) that ERT would provide if this condition is not satisfied. Note that this outcome is also recognised by AR2.

Finally, AR2 is questioning the significance of our novel upgrading approach as it addresses sites with specific characteristics. It is correct that we assessed this new methodology on one tabular three-layer structure. However, as explained in our manuscript, this structure was chosen because representing a typical regolith sequence of soil–saprock/saprolite–bedrock that mirrors the subsurface of many natural areas. The upgrading procedure was experienced on a specific conductive-resistive-conductive 3-layer sequence in order to mirror the Weierbach catchment case study setting. This approach may also work in other contexts, such as for example the reverse case, i.e. "resistive solum / conductive subsolum / resistive bedrock". Note that we especially invite the potentially interested readership for assessing the proposed methodology in this latter situation. Nevertheless, we agree with AR2 that the proposed upgrading approach is not a general improvement procedure as it is applicable at sites with a rather homogeneous top layer. However, we also want to stress that the proposed methodology is not "Weierbach catchment – specific". As discussed in our manuscript, the regolith of the Weierbach catchment is representative of the slate regolith which covers a large part of the Rhenish Massif. Hence, we anticipate that the proposed protocol could be used in manifold sites of this large central European geological area (extending from Luxembourg, through Belgium, France and Germany). Therefore, the proposed approach shall eventually be of interest for the hydrological sciences community working in this region. Moreover, as written in our manuscript, we further expect that our novel approach may also be transferable to other regions/catchments with similar characteristics, like forested catchments with similar bedrock geology.

<u>Specific comments</u>

My understanding of the proposed methodology is that it works on a site with a configuration of "conductive / resistive / conductive" three-layered structure, with a shallow layer "homogeneous in terms of resistivity and thickness." The authors propose to survey long profiles with large ES and smaller profiles with a smaller ES. And then, add some interpolated data points for shallow levels of apparent resistivity in the datasets of the long profiles. This is a bit tricky since it includes potential biases in the dataset of the long profile and poses questions in terms of spatial interpolation between smaller profiles and long profiles. Moreover, if you know that the shallow layer of your site is already homogeneous in terms of resistivity and thickness, what is the point to survey the site? And if it is actually not homogeneous, there is a great chance that this method will virtually tell you that it is, which is more problematic. It is also not clear what is the interpolation approach used in the method. Are the authors simply extracting a mean value for different depth levels of apparent resistivity and include that in the large ES datasets as a series of virtual quadrupoles along the profile with the corresponding depth level of apparent resistivity, or are the authors spatially interpolating the apparent resistivity of several small surveys with small ES into the large ES datasets? I think a diagram explaining the methodology would be highly beneficial for the reader to understand it a bit better.

From the specific comments above, we understand that on the basis of the current version of our contribution, any potential reader of our manuscript might miss a key step proper to the proposed upgrading procedure. The proposed methodology is indeed based on the incorporation of "virtual quadrupoles" defining several shallower levels of apparent resistivity. However, these virtual levels neither are constant mean values calculated for the entire study area nor result from a spatial interpolation between available shallow apparent measurements measured with a smaller ES. As explained in section 2.3 (step 3 of the protocol), ERT profiles using an oversized ES might be upgraded with surficial levels of apparent resistivity interpolated from a) calibrated relationships which were defined using a reduced number of shallow apparent resistivity measurements done with a smaller ES and b) the values of its own first acquisition level of apparent resistivity as input of the calibrated relationships. It is in this aspect that the innovation of the method lies. In the Weierbach catchment for instance, the four linear regressions resulting from the plot scale ERT profiles dataset done with an ES of 0.5 m (Figure 9) are intended to be used all over the catchment area. Thus, we plan to upgrade the ERT profiles measured with an ES of 2 m and covering all the catchment (white lines drawn in Figure 1) using as input the values of their own first acquisition levels of apparent resistivity (i.e. quadrupoles with external electrodes separation of 6 m). Note that we believe that the proposed upgrading procedure is well described in section 2.3, but we will try to clarify the text and/or introduce an additional diagram to better explain the procedure if requested.

AR2 mentions also that our novel procedure could introduce potential biases in the upgraded dataset. We are not exactly sure which bias AR2 is referring to. However, note that we already highlighted limitations of the proposed method, as well as potential improvements to face it, in the discussion section 4.2. For instance, we pointed out some specific local areas in the Weierbach catchment where the method would most probably lead to erroneous results by inducing false inverted surficial resistivity layers (e.g. riparian zone, where solum and subsolum have been eroded). It goes without saying that the upgrading protocol will not be applied to these locations. But we want to remind that our study clearly demonstrated that the application of the proposed upgrading method to the catchment-wide ERT survey dataset relying on an ES of 2 m will lead to an overall improvement of the inverted results accuracy. As shown in our work, if we do not apply this procedure, the lack of shallow apparent resistivity related to the oversized ES of 2 m data will for sure induce a much more important general bias.

AR2 finally asks why we want to survey a site if we already know that the shallow layer is rather homogeneous. We assume that AR2 refers to the solum of the Weierbach. As indicated in our manuscript, the solum itself is not the primary target of our catchment scale ERT survey. The goal is to inform on the spatial variability of the regolith as a whole: weathering state, potential relation with hydraulic properties, depth to bedrock, spatial organisation/connectivity, new insights on the substratum further deep.

Descriptions of the synthetic results are a bit confusing, especially in section 3.1.1 which describe a lot of figures and tables (both in the manuscript and in the supplementary materials), which elongates the reading of the manuscript a bit too much. Infact, I think there are too much scenarios and results to describe for a paper. I would rather suggest the author to focus on a couple of scenarios to simplify the text.

To shorten sections 3.1.1 and 3.1.2, we are willing to describe in the manuscript only the Wenner-Schlumberger array results if requested. Results obtained for the dipole-dipole array, which are similar to Wenner-Schlumberger ones, might only be presented in the supplementary materials. However, as proposed by AR2, we disagree to reduce the number of synthetic scenarios studied. It is indeed the diversity of these scenarios with varying resistivity and thickness contrasts that leads to a comprehensive understanding of the impact of the lack of shallow apparent resistivity data induced by the use of an oversized ES on inverted ERT image accuracy.