Response to Editor Ty P. A. Ferre for his comment in lieu of additional referees on "Imaging groundwater infiltration dynamics in karst vadose zone with long-term ERT monitoring" (manuscript hess-2017-477)

We wish to thank Editor Ty P. A. Ferre for his thoughtful comment that points several improvements to be done in the manuscript. For a better legibility, we have subdivided below the review comment (R) in several paragraphs to which we bring specific answers (A).

R: The authors have addressed a particularly difficult topic within hydrogeology: the dynamics of water flow in complex, layered, heterogeneous regions including relatively large preferential (karstic) pathways. They have applied an accepted and widely used hydrogeophysical method (electrical resistivity tomography) to this task. The strength of the study lies in the unique long-term (three year) data set in a karstic environment.
The challenge, not surprisingly given the complexity of the system, lies in the interpretation of hydrogeophysical data and the transferrability of the method and results. As is fitting for an exploratory application of a new method, the study was conducted in a very well characterized area. It would be worthwhile commenting on how well the ERT data could have been interpreted in the absence of this additional data. This has direct relevance to the use of the ERT method for other, less well-characterized sites. Additionally, the system under study is particularly amenable to study because the water table traverses the known conduits in a typical year, flooding them during winter and running dry in summer. This, again, indicates a wise choice of method for a specific study area. But, it would be worthwhile to comment on this specifically when suggesting that the ERT method could be useful at other karst sites. In other words, it is well known that ERT can only monitor dynamics in as much as it identifies temporal changes in water saturation. How can a reader decide if those conditions are likely to exist at measurable levels at a site before deciding to conduct an ERT survey?
A: The unique long-term data set is indeed one of the strengths of this study. One of the messages of the manuscript is however that no strong interpretation could have been drawn without characterizing the site in details, i.e. without the structural and lithological information gathered throughout this study. Although being well characterized, the site might not be ideal to conduct such ERT monitoring. The presence of a pronounced topography and highly dipping geological layers is likely to enhance run-off and infiltration processes that are complex to measure. The ideal site would indeed be a flat karst area with horizontal geological layers. In such cases, the occurrence of groundwater reservoirs within the epikarst would likely be easier to image with ERT monitoring. At the same time, despite the numerous challenges associated with the site under study, we could successfully investigate hydrological processes because of the amount of complementary information gathered throughout the study. In other words, we strongly suggest similar experiment to be conducted in areas where additional information could be collected, i.e. accessible cave systems, implementation of in-situ hydrological measurements, etc. It is also to say that the location of the ERT profile was not set randomly. As mentioned at the beginning of Sect. 5, it required seven preliminary ERT surveys over the study site to select an area with a large spatial distribution in terms of resistivity. This highlights the fact that such ERT monitoring experiment cannot be implemented anywhere. Preliminary studies are necessary to assess the feasibility of such a technique at a given site, and more specifically at the appropriate location within each site.

As for measurable levels of temporal changes in water saturation, they are strongly linked with the climatic conditions of each site. Seasonal rainfall averages are a good first indicator to know whether measurable changes in groundwater saturation are likely to occur at a particular site or not. Also, the geological layers being surveyed are important to take into account, which explains why preliminary ERT surveys are important to conduct at any site. We propose a few sentences at the beginning of Sect. 5 to highlight this aspect.

R: The ERT surveys appear to have been very well designed. The combined use of DD and GD surveys is thoughtful and the description of their differences in terms of spatial information and practical limitations is succinctly stated. The use of automated data collection and analysis, incorporating reciprocal measurements where available, gives confidence in the data quality. Similarly, the authors' recognition of temporal variations in contact resistance is noted as this is often overlooked in long term studies. I was impressed by the approach used to correct for temperature effects. However, it isn't clear to me that the heat transport was coupled with water flow. Given the complexity of the hydrologic system, this may not have been possible. But, it would be good to add more detail regarding how the temperature distribution was determined to allow for temperature correction. (This may be suited to an appendix.)
A: We actually removed the description of the temperature correction from a first draft of the paper, as we found that it added unnecessary details and length to the manuscript. But it will indeed be good to add such details as an appendix. As for the model used to calculate the temperature field, it was not possible to couple the heat transport with water flow. Water flows are what we want to identify with the ERT monitoring, which requires not fixing this parameter in the temperature model. In any cases, to our knowledge, there is no ERT monitoring study that uses a temperature correction that incorporates water flows in the solving of the heat transport. Similarly, the water saturation of each layer is not taken into account in the temperature correction of the resistivity, for the same reasons.

R: It is not clear to me if the resistivity inversion is spatio-temporal or if each time slice was inverted separately. Given that you are looking for changes in time, it seems that spatio-temporal inversion may be more appropriate. But, I would like to have seen some discussion of this choice. It also strikes me that interpreting each time slice independently may be more subject to imposing small scale seasonal variations in areas that are actually not seeing any real variation. It would be very interesting to see if a time lapse inversion results in as good a fit with some areas showing no seasonal EC changes.
A: As explained in Sect. 5.2.2 of the present manuscript, we use a time-lapse inversion procedure. This means that a time regularization constraint is used, linking each of the inverted models to the reference model. We will clarify this point in the manuscript to avoid misunderstanding.

R: In the end, I felt that the strongest element of this paper was the structural interpretation. This would be strengthened by more discussion of the process by which the arbitrary conductivity-bounds between regions were determined. It would be far more useful if this were explored automatically, perhaps using clustering techniques to propose alternative structural maps. For me, I think that some discussion of the EC limits is needed for publication. But, the paper would have more impact if this analysis were expanded and potentially seen as the basis for forming competing structural hypotheses. The danger as presented is that the authors may have unconsciously chosen EC limits to

confirm their pre-existing structural interpretation. This would, of course, limit the value of all of the work that went into collecting the data.

A: It might be important to stress that our main aim was to identify areas of contrasting behavior rather than delineating such areas (i.e. identifying the limits between them). Automatically detect boundaries and even time-lapse changes within ERT images is definitely an interesting topic. While clustering techniques could certainly bring additional sources of investigation, they also raise several questions about the way they should be implemented. For example, Xu et al. (2017, DOI: 10.1016/j.jappgeo.2017.07.006) has recently addressed clustering problematic within a short time-lapse ERT experiment at the Lascaux cave (France). Their analyses focus on the clustering of an ERT image in several clusters based on resistivity values of a single image. In their study, clustering is however not based on temporal changes within the ERT images. The risk of using clustering based on the resistivity values of a single ERT image is that it neglects the geometries and the dynamics of each cluster. Areas showing similar resistivity would be associated, while they could reflect layers of different lithology, saturation and/or clay ratio. Performing hierarchical agglomerative clustering on the entire time series seems therefore more relevant. Such an approach could for example focus on the clustering of the correlation matrix of all the cells of the resistivity image for all the time steps. Such a method seems interesting but certainly requires further analyses and synthetic modeling to select the best parameters e.g. to pick the best method for calculating the distance between clusters. We believe that this goes further the scope of our paper but we propose to add a few words on these aspects as perspectives in the conclusion.

Furthermore, such a technique would not take into account the geometries of the sub-regions within the ERT image. In our case, the highly dipping conductive feature present in the ERT survey is likely to be associated with other, less dipping conductive layers, which does not make sense from a structural point of view. In other words, we believe that our site might not be ideal to start investigating the use of automatic clustering tools, as being too complex in terms of geological structures.

This is the main reason why we originally proposed to subdivide our ERT image in 8 sub-regions “based on their average resistivity values and arbitrary thresholds”, as explained in the manuscript (Line 557). To clarify this, we propose to add a table with statistical analysis of each cluster (i.e. mean, median, standard deviation through time, etc). We are aware that such an approach is less transportable to other case studies but will contribute to clarify our choice of different clusters

R: Unfortunately, and not unexpectedly given the complexity of the system under study, I found that the hydrologic interpretations were somewhat qualitative. It is interesting to see that there are correlations and delays between responses. But, it doesn’t seem to rise to the level of increased understanding of flow dynamics. This may simply be a matter of emphasis - you could highlight what was learned from the ERT that would not have been possible without it. But, it reads more like using your hydrologic insight to give plausible explanations of what you see in the ERT results. Understand, this isn’t a strong criticism. I think that this is an advance and shows potential future use of ERT for monitoring dynamics in some karstic systems. But, I think that it is a mistake to make this the emphasis of the paper - starting with the title. Rather, I would focus the paper on the ’hydrostratigraphic’ results - showing how you could use the time lapse ERT to identify structure in the subsurface. That could be expanded and extended and then it would be appropriate to say that this interpretation is consistent with what was seen in other hydrologic measurements. As an added benefit, this would allow you to shrink the hydrologic section, which is not as tightly written as the previous sections.

A: We appreciate this suggestion. The current version of the manuscript could indeed be reworked to emphasize the ‘hydrostratigraphic’ aspects. We therefore propose to

rearrange the manuscript as suggested, firstly focusing on the ERT results, and then on the structural and hydrological data. This could indeed highlight what could not be learned without the ERT results. However, the interpretation of the ERT dynamics in terms of karst hydrology is for us one of the important aspects of the paper. Especially, the joined analysis between time-lapse ERT results and percolating water measurements is definitely a novel approach that is promising to investigate the sources of distinct in-cave flow types and their lithological/structural constraints. As already explained in our response to Referee #1, we are working on a future paper in which we will focus on a lumped karst modeling of the vadose zone infiltration processes based on the drip discharge data (using the KarstMod modeling platform, from Mazzilli et al., 2017, DOI: 10.1016/j.envsoft.2017.03.015), and their relationships to the ERT data. However, this could not be done without highlighting the role of the lithology and the structures in the time-lapse resistivity changes. In other words, we see this present manuscript as an original study in which the sources of distinct percolation types are imaged and further linked to geological structures of the karst system under study. We believe that rearranging the section of the manuscript as proposed by Editor Ty P. A. Ferre could clarify our message.

R: In summary, I think that this is a very strong paper and that it should be published in HESS. But, I think that the current emphasis on flow dynamics is not ideal. Rather, it could be a very novel and interesting example of using dynamics to better define structure. This would be most interesting if it could be done automatically, e.g. using clustering tools, and if it led to multiple competing hypotheses that could be further tested in the field.

A: Using ERT dynamics within clusters to better define structures would definitely be impressive. But we believe that this goes further the scope of this paper as this would require numerous synthetic modeling for calibrating a working clustering procedure. Moreover, as already pointed out, the site under study might not be ideal to apply and validate such workflows, especially given the complex geological features. Our study however paves the way for such techniques to be tested. Therefore, to strengthen the message of our manuscript, we propose to rearrange the different sections to highlight what was learned from the ERT that would not have been possible without it, as suggested above; the main goal of the paper being to investigate the link between underground structures and percolating water via this long-term ERT monitoring.