We are grateful Dr. Niklas Linde for his careful comments, corrections and suggestions. His valuable review helped us much to improve our manuscript.

The manuscript of Tran et al. address important questions that needs to be addressed for efficient and reliable integrated terrestrial Earth monitoring and interpretations using geophysical data. They couple a state-of-the-art two-phase (gas and liquid) and three-component (air, water, heat) simulator and its associated inverse capabilities with an electrical resistivity tomography (ERT) forward solver. By inferring the values of the petrophysical parameters they can use the ERT data together with point measurements of state variables to infer various properties of the subsurface. A synthetic test example shows that ignoring the influence of temperature on the resulting electrical conductivity leads to biased estimates of hydrodynamic properties. They also present an application to the Rifle site, Colorado, and show that the method is capable of parameterizing a subsurface model such that it can make quantitative predictions of various state variables. What I like with this paper is that it highlights both the need and the challenges of fully-coupled hydrogeophysical inversion. Clearly, multi-phase and several components must be included. This leads to many unknowns to estimate and the estimation is quite difficult to achieve due to non-linearity and the limited information content in the available data. Also, I do appreciate the attempts of using a global sensitivity analysis to enable the parameter estimation to focus on the most important parameters. I judge that this would be a very nice contribution to HESS after moderate revisions.

Comments: 1. It is a bit surprising that variations in solute concentrations are not considered. One would expect that the rainwater is much more resistive than older formation water. I realize that this is challenging, but mentioning this in the discussion as an outlook might be appropriate. Also, there is no information (at least I did not see it) about what salinity/electrical conductivity of the pore water that was assumed. I assume that a constant value was chosen. Was it based on actual water samples or was it a fitting parameter to get the right amplitudes in the resistivity model? This needs to be clarified.

Reply: In our case study in Rifle, the electrical conductivity of pore water does not much change in time and space. The measured water electrical conductivity ranges from 0.237 to 0.255 S/m at two different wells over the simulation period. Therefore, we selected the average value of these samples σ = 0.244 S/m. We presented this information in line 12-13, page 8 as below:

"The electrical conductivity of water, which does not vary significantly over time at the Rifle site, was taken from the measurements at the nearby wells, and is equal to 0.244 S/m."

We also discussed the approach that considers the solute dynamics in hydrogeophysical inversion in the revised version (lines 13-16 page 8) as below:

"In case that the spatio-temporal variation of solute concentration and resulting electrical conductivity in pore water are significant, it dynamics should be simulated by considering it as a component in TOUGH2. A formula that links the solute concentration

with the water electrical conductivity is also needed to develop.”

2. The hydrological model contains simply two geological layers with uniform properties. Clearly, the ERT response must consider the deeper (aquitard) Wasatch formation. I assume that the authors do this, but again, this is not stated in the manuscript. In short, the model domain is different for the electrical problem and this must be clarified. Perhaps the authors embed the simulated resistivity in the fill and alluvial layer into the model shown in Figure 3a?

Reply: We discuss how to map the apparent resistivity from hydrological to geophysical computational mesh in the revised version (lines 5-12 page 14) as below:

“The BERT computational mesh was automatically generated in BERT. We set the maximum cell size, which controls the mesh refinement at a small value (0.2 m) to capture the local variation of the soil electrical resistivity. Other parameters that determine the BERT computational mesh were kept as their default values. For more information about BERT mesh generation, we refer to Rücker et al. (2006). The apparent resistivity was mapped from the hydrological to the BERT mesh using the nearest distance method, i.e., a cell in the BERT mesh will get the resistivity value of its nearest cell in the hydrological mesh. The electrical resistivity of the Wasatch layer was set at its average value obtained from geophysical inversion ($\rho_b^{Wasatch} = 45 \ \Omega.m$).”

3. The only lateral heterogeneity in the model is due to a spatially varying interface between the alluvial fill and the alluvial layer. Clearly, any ERT delineation of this interface will be highly uncertain (especially as the water table divides the gravelly sand into a saturated and an unsaturated part. I assume that the interface was guided by an observed interface in well TT02? This should be explained and the sensitivity to errors in this interface should be discussed in the paper.

Reply: Yes, we selected the threshold of electrical resistivity for determining the interface between layers using observation at TTO2 and TTO3 wells as references. We added this information in the revised version (lines 15-22 page 13) as below:

To specify the locations of the fill-alluvium and alluvium-Wasatch interfaces from ERT geophysical inversion, we used the depths of these interfaces observed at the TTO2 and TTO3 well as the references to determine resistivity thresholds. Accordingly, a grid cell with a resistivity greater than 1.52 Log_{10} ($\Omega.m$) and above 1.5 m depth belongs to the alluvium layer. The cells whose resistivity values are smaller than 1.83 Log_{10} ($\Omega.m$) and below 5 m are assigned to the Wasatch layer. The remaining cells are in the fill layers. The magenta and white lines in Figure 3 represent the fill-alluvium and alluvium-Wasatch interfaces, respectively.

Minor comments: Page 1, Lines 18-19: Are the measured data relatively accurate and the simulated data reproduce them (what is written) or is it the predictions that are relatively

good (what I assume the authors want to write). Please correct. Also, try to be quantitative and replace "relatively" with quantitative measures.

Reply: The sentence was reformulated (lines 18-19 page 1) as below:

"At the Rifle site, the coupled, hydrological-thermal-geophysical inversion approach well predicted the matric potential, temperature, and apparent resistivity with the Nash-Sutcliffe efficiency criterion greater than 0.92."

Page 2, line 4: Perhaps "play an important role in controlling" instead of "control". Other factors control this partitioning as well.

Reply: We changed this sentence to "*For example, watershed moisture content and temperature are the main factors that control the partitioning of precipitation into evapotranspiration, infiltration and runoff*" (lines 3-4 page 2).

Page 2, line 5: Is ecosystem moisture an accepted term?

Reply: We changed to "*For ecosystem, moisture content and temperature conditions*" (line 6 page 2)

Page 2, line 10: Perhaps "spatio-temporal". Reconsider the term "native". Generally speaking, this sentence would be easier to follow by clarifying what are the differences in scales between the so-called native processes and ecosystem functioning.

Reply: We changed to "spatio-temporal". The term "native" was replaced by "local". This replacement should clarify the scale differences.

Page 2, line 12: I suggest to remove "direct measurements". Soil moisture is typically inferred from dielectric properties. This is not a direct measurement. Indeed, one can discuss if direct measurements exists altogether.

Reply: "direct measurements" was removed

Page 2, lines 21-24: Data and models are not the same! Hubbard et al. (2001) integrate geophysical models/tomograms with point measurements. On line 24, add "pumping tests" as these are needed to get the average hydraulic conductivity, while the flowmeter gives relative variations.

Reply: We restated the sentence (line 20-22 page 2) as below: *"Statistical approaches have been extensively used to integrate point measurements with commonly geophysical models/tomograms, such as Ground Penetrating Radar (GPR) and Electrical Resistance Tomography (ERT)."*

"pumping tests" was added in the revised version

Page 3, line 2: I think it should be "hillslope".

Reply: This was corrected

Page 3, lines 3-4: Unfortunately, few hydrogeophysical inversion approaches are developed to "improve quantification of subsurface processes". They are in most case only used for parameter estimation, which is something very different and much more restrictive.

Reply: We changed the "improve quantification of subsurface processes" to "estimate soil hydrological parameters". (lines 4-5 page 3)

Page 3, line 6: You need a qualifier before inversion, something like "hydrological".

Reply: We modified the sentence (lines 6-8 page 3) as below *"Because coupled inversion approach links hydrological simulation outputs (e.g., moisture content) with geophysical data, it avoids the errors typically associated with geophysical inversion process"*

Page 3, line 13: I don't think that "its" is used correctly here. Please rephrase.

Reply: We changed "its" to "their"

Page 3, line 24: "at high resolution (add references)".

Reply: We deleted the paragraph that contains this phase

Page 4, line 1: This is true for near-surface applications (please clarify). Seismological stations are probably used much more for geophysics in general. I would call ERT a method, not an approach.

We modified the sentence in the revised version (lines 21-22 page 3) as below: *"To date, ERT is the geophysical technique that is most commonly collected in an autonomous manner for near-surface applications"*

Page 4, line 12: Replace "and" with "which". Line 17, remove "the", line 18 change to "images". Line 26, replace "into a" with "within".

Reply: These were changed.

Page 5, line 7: I would remove references to Archie here. Archie defined the formation factor and the effect of saturation. However, referring to Archie implies that surface conductivity is ignored, which is not done here.

Reply: "Archie" was removed

Page 6, line 8: Change symbol for porosity to be consistent with the rest of the paper.

Reply: The symbol was changed

Page 6 and elsewhere: Equations are part of the sentences, so use "," and "." As appropriate after the equations. Ensure that variables are in italics. At the moment, they are sometimes, but not always (e.g., line 19).

Reply: Commas were added after the equations

Page 7, line 18: Perhaps extend this aspect in the discussion? Is this a big issue?

Reply: In the Rifle site study, transpiration and plant water uptake are not a big issue in this study because the study site is mostly bare soil. However, in the case that soil is largely covered by vegetation, vegetation should be considered. We are working with a land community model (CLM) to address this issue.

Page 8, line 11: Why not use subscript "b", as sigma_a or rho_a suggest apparent properties and not bulk properties.

Reply: We replaced subscript "a" by "b"

Page 8, lines 17-18: It is ok to fix the cementation index, but I am not sure if I see the argument here. The surface contribution is a constant over time while the saturation contribution is time-variable, so identifiability should not be an issue if enough variations in the states are probed.

Reply: Our sample measurements show that water conductivity does not much change over time and space. Because we had measurement value so we fixed in the inversion. We don't know surface conductance of the fill layer so we attempted to estimate it.

Page 9, lines 24-29: How did you balance the different types of data? Ideally, one should have a WRMSE of 1 for each data type after inversion and using actual data errors. In practice, this is essentially never done. The ERT data are given errors of 5% to somehow accommodate modeling/petrophysical errors, but they are presented as observational errors while the actual observational errors are probably less than 1%. I assume similar common "tricks" are used for the other data types. I don't criticize this, but some discussion about how the data were weighted and reporting the final data misfits for the individual data types should be part of the paper.

Reply: We used measurement errors (inverse covariance matrix in equation 15) to account for different types. It is reasonable because in theory, it represents the likelihood function. In application, it implies that the more accurate observation will have stronger weight in the objective function. The "trick" here is that we did not know exact measurement errors, so we subjectively set measurement errors based on our evaluation of the accuracy of each measurement type.

Page 12, line 3: "at the within". Please revise.

Reply: This was corrected as below (lines 3-4 page 12): "The newly developed approach was tested at a floodplain adjoining the Colorado River, near Rifle, Colorado"

Page 13, line 13: It should be Günther. Also Rücker when his name is used.

Reply: This was corrected

Pager 13, lines 15-18. This is a bit of a circular argument (a similar construct is also found later in the paper). Why not write something like: As expected, the clay-rich fill and Wasatch layers show up as less resistive layers in the ERT : : :".

Reply: We reformulated these sentences to "As expected, the clay-rich fill and Wasatch layers exhibit less resistive than the alluvium layer" (line 8 page 13).

Page 13, line 18: Why not across the full ERT line?

Reply: In our project, we need to provide information on hydrological and thermal dynamics for other studies at a part of the transect around the TTO2 well. So to save the computational time, we performed inversion in this part of transect only.

Page 13, lines 21-22. Give references that support these values.

Reply: That is the average value estimated from our core drilling and sampling.

Page 13, line 23: This is ok for the hydrological model, but clearly insufficient for the ERT modeling. Express how you extend the ERT model and account for the Wasatch layer.

Reply: See our previous response.

Page 14, line 2: "the meteorological".

Reply: This word was corrected.

Page 14, line 13: Write "apparent resistivity".

Reply: This phrase was corrected.

Page 14, line 14: State what this implies in terms of assumptions. Using data from a well that is not located in the modeled study area.

Reply: We add the following assumption to the revised version (lines 27-28 page 14):

Assuming that the lateral variation in subsurface temperature between TTO2 and TTO3 wells (see the TT03 location in Figure 2) was insignificant, we used the temperature data at the TTO3 well for inversion.

Everywhere: Please consider to replace "heat conductivity" with the more common term: "thermal conductivity".

Reply: We replaced "heat conductivity" by "thermal conductivity"

Page 18, line 14: Replace "ignorable" with "negligible".
Reply: We replaced "ignorable" with "negligible".

Page 19, line 5: It should be "reproduce".

Reply: It was corrected

Page 19, line 27: Write "high apparent resistivity values", same for line 28. Also, on line 28: replace "located deeper" with "more sensitive to deeper locations".

Reply: These were corrected

Page 20, lines 1-5: This problem would not appear in a log-log plot and it would be more consistent with the error model.

Reply: In these sentences, we meant that because we used relative error to set the measurement error of ERT. As a result, the absolute error of higher apparent resistivity is larger than that of lower apparent resistivity. This leads to lower weight of high apparent resistivity in the objective function.

Page 20, line 21: It seems impossible that the downward flux is 24 mm/s. Please clarify what this value refers to? I interpret this in the same way as a Darcy flux and this seems unrealistically high.

Reply: We corrected this information. It it 24 mm/day.

Page 20, line 25: Replace "to the upward flow up to" with "upward flow starting at".

Reply: We replaced "to the upward flow up to" with "upward flow starting at".

Page 21, line 5: Add "estimated" before "subsurface".
Reply: We added "estimated"

Page 21, line 15: State how much noise was added.

Reply: We modified that sentence as below (lines 4-5 page 22):
"Add Gaussian white noise (zero mean and standard deviation of 5% of artificial

apparent resistivity data) to the artificial data to obtain the synthetic data."

Page 22, lines 22-24: This is a worthy contribution, but it seems as if not accounting for the variations of salinity is a bit limiting. The precipitation will clearly be much more resistive than the groundwater. Perhaps something to comment on?

Reply: We agree with the reviewer that considering the variation of salinity is necessary. It would be implemented in the next study. We added a sentence that discusses how to consider this variation in hydrogeophysical inversion in line 13-16 page 8.

Page 24, lines 1-4: Perhaps comment on the potential of filtering approaches/data assimilation in the context of this type of problems?

Reply: We add some filter techniques (Kalman ensemble filter, maximum likelihood ensemble filter, particle filter) that we will use in future research (line 19-20 page 24).

Figure 2: What are the East-West trending yellow dots. What are the green dots. Explain or remove.

Reply: These dots show the locations of other biochemical measurements. We removed them

Figure 3: The depth is positive with depth, so please remove "-" signs". Also give units and properties shown (log10 of resistivity). Add "inferred" before "fill-alluvium". Also, explain how this was done somewhere in the paper.

Reply: The figure was corrected. The explanation that shows how to determine the interfaces between layers was added to the revised version.