Author Response to Reviewer #2:
Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-559, 2016.

First and foremost, we would like to thank the reviewer for donating their time to review our manuscript and for providing a fair and constructive review of our work. These comments and suggestions will undoubtedly improve the impact and utility of our paper.

Reviewer 2: General Comments

This manuscript is on the topic of geophysical and traditional measurements of a reach of river to investigate the suitability for time lapse ERT to study river bottom processes. The writing is in clear, good English, and the figures are mostly very readable and nicely drafted. The topic – either from the Hydrology or Geophysics perspective – certainly has the potential to be of interest to HESS readership. I believe the topic of this work fits into the scope of this journal. The most significant limitation I see to this work is related to the experimental design, which is largely absent from the writing. In short, it is difficult to tell what was being tested about the hydrology, and why measurements were implemented to carry out that test. The stated hypothesis is apparently related to "will the geophysics work," while the theme of riverbed processes appears and disappears throughout the manuscript. In the end, I remained confused about exactly what the reader was meant to take away from this given the setup of the writing and the design of the study. There is certainly lots of good data here and on some level this has the potential to be of high interest to the hydrology community, but there is a need for substantial revision for focus.

Response: We believe the reviewer has provided a fair assessment. The problem statements and hypotheses we set out to test could be described more effectively in the introduction; Reviewer 1 raised a similar comment. Because there were no previous examples of ERT being used to investigate riverbed dynamics in fractured rock, this study was in part, a type of proof of concept. We wanted to test the utility of ERT in a bedrock river environment under natural field conditions. However, the motivation for this work was directly associated with furthering our understanding of potential mechanisms associated with a groundwater-surface water interaction in a fractured riverbed system. That being said, we believe we can significantly improve the introduction and associated text throughout the manuscript to ensure a clearer "takeaway message".

There were several other notable issues/limitations related to measurement methods, data processing, and absence of some measurements that are detailed in my General and Line-by-line comments below. At this time I am recommending this manuscript be returned for major revision, however if the experimental design is not substantially clarified and the focus reworked to highlight hydrological interpretations, a second review would likely not result in a favorable recommendation.

There is a substantial disconnect between the topic of the science question and the posed hypotheses. Although the science question is not explicitly stated, it is my best interpretation that the following reflects the intent of inquiry: ". . . there remain gaps in our conceptual understanding of groundwater-surface water interaction and exchange mechanisms in bedrock rivers where discrete fracture networks will dominate groundwater-surface water flux with secondary interactions supported by the porous rock matrix." On the other hand, the hypothesis is explicitly stated, although it appears to be limited to a yes-or-no "will it work" type of speculation: "we hypothesize that a groundwater-surface water mixing zone – encompassing fracture and matrix flow and diffusion – may be identified within a fractured bedrock riverbed by monitoring spatiotemporal changes groundwater temperature and porewater electrical conductivity using minimally invasive electrical resistivity methods." Further complicating matters is the text between Line 69 and 76 that highlight the hydrological outcomes while disregarding the stated hypothesis.

Response: We appreciate the reviewers concerns and believe they can be addressed through reorganization and emphasis of key contributions.

Throughout the manuscript, speculative statements about river ice, river-bottom ice and frost are made, though they do not appear to be supported by any direct measurements or observations. Estimates of loss along reaches based on calculations of discharge using rating curves in conjunction with stage height monitoring appear to be absent. This line of evidence would substantially help to support geophysical observations.

Response: As stated in the manuscript the winter freeze-up period was accompanied by many challenges. In hindsight, a very different approach might have been used to fully understand the impacts of ice on the measurements, but we had not anticipated the conditions that we experienced. There was little that we could do to directly assess the ice in the river given the resources we had at our disposal. This is acknowledged in the paper. For this reason, the river ice is examined in the discussion rather than the results to avoid confusion between what was measured and what is interpreted. This paper focuses on the geoelectrical dynamics within the riverbed; we do not attempt to provide analysis of larger-scale hydrologic flow system. Estimates of loss/gain were measured at this site but are being prepared in a separate hydrological paper by other workers. Future work may integrate these data sets but at this time we remain focused on the geophysical transience and processes occurring with the upper few meter of rock.

I felt that the following questions posed early on in the manuscript were not clearly answered: Do you find that groundwater-surface water interaction was restricted by poor vertical connectivity and limited bedrock incision? Did you find that groundwater-surface water connectivity through discrete fractures was highly variable in space and time, and depended on fracture size or aperture, river stage, and the distribution of hydraulic head within the flow system?

Response: We agree with the reviewer. We will provide a clearer set of conclusions based on the questions posed in a revised introduction. A similar comment was raised by Reviewer 1.

There is a huge amount of data contained in this manuscript, however in some cases data was left unused in interpretations and discussion. For example, precipitation & snowfall, daily river stage, fracture content as a function of depth, atmospheric temperature, etc. Why include these data if they are not utilized? In the end, if the hypothesis was to test “will ERT work for this” I think that was not clearly answered, and furthermore, given the high dependence on temperature, it may be that that answer is “no.”

Response: These data provide critical context for the reader; they also support elements of the conceptual model (e.g., fracture networks in the rock). The hydrology data summarized in Figure 4 provides critical context for the geophysical observations and discussion of seasonal trends. Our discussion and interpretation of seasonal transients in the geophysical measurements implicitly utilizes the hydrological information. Without these data the reader would not be able to assess the severity of the seasonal temperatures fluctuations and river stage (presented here as flow), precipitation and snowfall on the geophysical response, or appreciate the frequency of our measurements within spectrum of field conditions. These data serve as a critical point of reference or comparison for other investigations in rivers. Further, the exclusion of these data would not improve the manuscript nor would it reduce the number of figures, thus, we support the inclusion of these data.

Nevertheless, the specific contributions of this study to the conceptual model will be more clearly stated. One of those contributions is the strong influence of temperature, which is viewed as an indicator of the vertical extent of surface water influence (direct or indirect) within the bedrock riverbed.

Specific Comments:

Line 55: Perhaps add a comment on what Fan et al., 2007 found here?

Response: Clarifications will be implemented.

Line 63: There is a lot going on in the figure and it is weakly linked to the text. Are you testing these concepts?

Response: The conceptual model is meant to represent the system in its entirety. We agree that the figure was not as strongly linked to the text as it should have been. We believe this can be addressed. It is important to present the conceptual model and elements that, in theory, could be explored or tested with the approach used in this study. There are, however, limitations to our study. The figure serves as motivation or starting point to our conceptual understanding of groundwater-surface water interaction or hyporheic zones in fractured rock. It is a simplified view of a fractured rock system, showing advection and diffusion processes for a gaining and losing stream; these are well-known concepts in sedimentary fractured rock. We do believe that the text can be improved to better link our study to the elements of the conceptual model.

Line 92: The Singha paper has 2014 printed on top of it, but I'm not sure which date is correct.

Response: It appears to be published on-line in 2014. But the paper wasn't fully published until 2015.

Line 175 – 180: Was the formation of basal ice actually observed at the site or only inferred?

Response: Basal ice was visually observed in the field as indicated in the manuscript. The ice was no longer visible once the river froze over.

Line 213 – 216: I am unfamiliar with this method of sampling temperature while the sensor is in motion. Certainly the sensor itself, however small it might be, has some thermal mass that requires time to equilibrate to the surrounding water temperature. Even though the sensor is capable of measuring at 0.5 s rate, that does not mean that the measured data are reflecting changes in the formation at that rate. A reliable reference should be included here to justify the method, and a controlled validation test and sensor calibration under laboratory conditions should be conducted to quantify sensor response.

Response: This is very interesting question and one that we had not considered until now. Based on the information we received from the RBR*solo*™ manufacturer (RBR Limited, Ottawa, Canada), these sensors will resolve 63% of a full-scale temperature change in 1 s, 95% in 1.5 s and 99% in 2 s. This particular sensor has a maximum resolution of 0.00005°C (full scale). However, in this study we only report temperature changes to the 0.01°C. Given our reported rate of decent ~0.8 cm/s combined with our temperature resolution, a conservative vertical "averaging" estimate might be ~1.5 cm based on the full resolution. Therefore, sensor response time in this case appears to be very small and negligible on the data sets presented. Unfortunately, we do not have the laboratory equipment to make any further comments on the performance of this sensor deployed in this way.

We can include some of this information in a revised manuscript.

Line 219: What is the value of measuring snowfall accumulation if snow density is not also reported with a conversion to SWE?

Response: Snow accumulation is reported as SWE. Clarification of SWE will be provided in the figure caption.

Line 221/Figure 4: What scale are the red dot “Resistivity Samples” on? They appear to be only temporal and unitless, however they seem to track the river stage which is confusing. Is ‘snowfall” in this figure converted to SWE? If not, please do and clarify the label.

Response: The resistive samples “red dots” are plotted to show their temporal position and sampling interval during the seasonal hydrological conditions of the river. They track the river stage (y-axis) because these measurements were collected in the river and the stage is explicitly used in the models. These red dots effectively identify the stage conditions sampled in this study.

Line 228: “effective sensing depth” does this mean measurements are reflective of the 6m depth zone, or the entire aggregated zone 0 to 6m depth?

Response: The instruments sampling depth (volume) is defined by the impulse response function (McNeil, 1980). Here, the sampling depth is stated as 6 m which is general rule-of-thumb to the depth of investigation for the instrument in this orientation. More descriptive phrasing can be added to the text.

Line 237: “BLANKED by bedrock rubble” I am not familiar with this usage of “blanked” in this context. Suggest rewording for clarity.

Response: The word was “blanketed” but we can change this to “covered” to avoid confusion.

Line 239 – 244: Does the electrode construction method have any particular importance to this study? This sounds like very typical ERT cable construction, albeit by the end-user rather then a professional fabricator. Probably could be deleted.

Response: In theory the construction of our cables is similar to commercial systems, but because it isn’t a commercially available cable we felt it best include the details of its construction. There are design elements that could have an impact on the results (e.g., electrode construction and length) that a reader might want to know. Our inclusion of the design of our cable is consistent with the approach of other workers (e.g., Van Dam et al. 2014).

Line 261: How was the measurement time determined? It is known that diurnal fluctuations in stream water temperature may be of magnitude in excess of 10C, similar to your annual range of groundwater temperatures. Also you acknowledge the affect of temperature on the ERT readings; how does the timing of the measurements affect the data due to daily fluctuations?

Response: This is a great question. Although we discuss the limitations of this study (data aliasing) and potential sources of thermal influence on the geophysical measurements, we do not neglect temperature fluctuations. These are considered to be an important component of the observed geophysical dynamics. Our sampling frequency considers longer-period (seasonal) variations rather than diurnal variations. Given our coarser sampling interval we cannot comment on the impacts of shorter-period temperature fluctuations (diurnal) on the groundwater resistivity. We discuss the potential impact of sunlight (heating) of the riverbed and its spatial/temporal variability on the geophysical signatures in the last paragraph of Section 5.1.

Further clarification on this subject can be incorporated as well as the inclusion of the references provided by the reviewer below.

Constantz, Jim, Carole L. Thomas, and Gary Zellweger. "Influence of diurnal variations in stream temperature on streamflow loss and groundwater recharge." Water resources research 30.12 (1994): 3253-3264.

Constantz, Jim. "Interaction between stream temperature, streamflow, and groundwater exchanges in alpine streams." Water resources research 34.7 (1998): 1609-1615.

Line 262: "manually filtered" What criteria was used for manually filtering? Why was this approach used rather than the common quantitative method of envelope filtering based on an error model?

Response: In our case, only obvious data outliers were removed (e.g., failed measurement based on a non-zero standard deviation); we intentionally did not apply any pre-inversion data smoothing or averaging in an attempt to preserve the data trends and maintain data-input consistency. However, data smoothing was direct applied in the inverse routine as described in the text. The approach used in study will depend on the site conditions and desired outcome of the experiment. In our case, we were concerned with preserving the signal of the natural system (governed by multiple factors) rather than enhancing a particular element or physical processes in the model. Additional details of the inversion setup will be added to the manuscript.

Slater, Lee, et al. "Cross-hole electrical imaging of a controlled saline tracer injection." Journal of applied geophysics 44.2 (2000): 85-102.

Line 264: "moderate to high damping" – Does this mean different damping factors were used on each dataset? What is the numerical value of damping used and how is this value incorporated into your inversion scheme?

Response: The initial damping factor was moderate but allowed to vary in the inverse routine; the same starting parameters were used for each model run. However, models were allowed to optimize the dampening factor depending on the model convergence. Therefore, the dampening factor likely increased with noisier datasets. Additional details including specific parameters used in the routine can be added to the text.

Line 267: What parameters, how were they optimized, and were identical settings used for all datasets?

Response: Additional details and clarification will be incorporated into the text.

Line 268 – 269: Certainly achieving the lowest possible RMS is not the optimal approach to achieve the most "believable" geophysical result. How does the RMS relate to observed measurement noise/errors? At what point is the inversion fitting noise?

Response: We agreed. That being said, minimizing RMS error (minimization between measured and modeled data) is a reasonable approach to achieving the most representative model. Of course, the model is only as good as the measurements inputted into the model. One of the challenges in this study was finding a reasonable solution to the data collected in the winter (frozen) months. These periods were accompanied by erroneous data points (largely due to the high contact resistances of frozen ground and ice) higher overall noise, and thus, model convergence and stability was at times challenging. The issue could only be circumvented by imposing a maximum limit on modelled resistivity. This meant that the model RMS errors would be considerably higher (note the data presented in Figure 10). Therefore, minimizing the RMS errors was not the only criteria used to achieve the most realistic model.

Line 280: Is there a reference for this Resistivity Index? What is the justification for manipulating the data in this way?

Response: The resistivity index (a broadly and routinely used normalization technique) was used here so that we could compare the transience observed between the pool and riffle sections, which had very different mean resistivities. The index simply allowed us to normalize the data sets to the mean value, permitting easier comparison of the timing and magnitude of transients at each site. Our approach is defined in Equation 3. We can also provide a better justification for its use in the methods section.

Line 293 – 294: Where is the data demonstrating upward head shown?

Response: The data was not included because it was not used in the study other than for the purpose of establishing the direction of potential groundwater flow. Since the data did not reveal any major changes in the gradient we decided not to include the data, and instead, simply state (in words) what the data showed. The hydraulic head data could be added to Figure 7.

Line 307/Figure 6b: It would be helpful to grade the colors of the lines linearly to more easily show the temperature trend. Even better would be to present these data as a matrix/grid where time is on the x-axis, depth is on the y-axis, and color represents temperature.

Response: This is a good recommendation and one that we will consider in the future. However, we think the current presentation is also reasonable and maybe more “understandable” at a quick glance. The purpose of the figure was to show the temperature swings (dynamics) with respect to depth, and the extent of the heterothermic zone, and illustrate the temporal variability overserved during the winter months. We think the current figure layout achieves these objectives.

Line 308: “correspond to areas” I cannot tell from the figure how the fracture patterns correspond to the temperature results. Perhaps some annotation, or another approach to presenting these data would help.

Response: We agree that some annotation to the Figure 6a should be added to better illustrate the position of fracture zones.

Line 321 – 331: [Figure 8] This seems more like a discussion point rather than a result.

Response: We agree. This sentence would be better placed in the discussion or conclusion section.

Line 356 – 358: “greater number of measurements. . .” why would the number of removed data cause higher RMS? Presumably if the data were removed, they would no longer be included in the RMS calculation.

Response: The missing qualifier in our text is that while obvious outliers were removed, the overall dataset remained nosier compared to unfrozen periods. Further explanation is provided above (see comments for Line 268 – 269), and necessary revisions to the text can be made.

Line 374: What are the observed thicknesses of basal ice and floating ice?

Response: Unfortunately we were not able to measure the thickness of basal ice. It was slimily visually noted in the field.

Line 415-416: “groundwater discharge in this section” I don’t follow the logic why the relationship between substrate resistivity and “surface water response” indicates magnitude of discharge that could be interpreted in this way. Also, correlation is not demonstrated or quantified.

Response: Here, we are simply saying the transience observed in surface water (days 13, 26, 31; Figure 7) are not readily apparent in the riverbed, and thus, surface water may not be interacting with groundwater during these periods. This suggests that groundwater discharge may be a more dominant/overriding process. We can modify the text to improve the clarity of this statement.

In this case, “correlation” is not accompanied by a statistical qualifier; therefore, it should be read as qualitative description of a relationship between to things.

Line 432: “strong upward hydraulic gradients” please indicate where this is demonstrated by data.

Response: There are two options: we can include these data in Figure 7 or simply state the gradient range in the text. Given the limited use of the data we will include the calculated gradient range in the main body.

Line 436: “likely dominated the bulk electrical response” Why ‘likely’? Based on the evidence shown, temp is clearly dominating the ERT signal.

Response: Agreed. This can be changed.

Line 445: Where is ground frost or riverbed ice formation measured data shown?

Response: Neither ground frost or river bed ice were explicitly measured in this study. Ground frost was interpreted based on the resistivity data (Figure 11e), while riverbed ice (basal ice) was observed in the field as described in the text.

Line 459 – 473: As previously stated, data showing the frost and ice should be shown.

Response: Please see response above.

Line 474: I’m not sure what evidence directly supports this statement. The provided sensitivity analysis appears to only vary the river water electrical properties; this doesn’t seem to directly simulate the presence of ice as claimed in this statement.

Response: The reviewer comments are understandable. We did not mean to suggest the “sensitivity analysis” was done to evaluate the influence of river ice. We can address this issue with some reorganisation of the text.

Line 479 – 480: Quantify this? Why would inputting a one-half of true river water resistivity lead to “substantial overestimates of river resistivity” – wouldn’t the river water resistivity be fixed so that the output = the input?

Response: We were referring to the implications of fixing the model with an inaccurate water resistivity. The point of the statement is to highlight the importance of accurately representing the geometry and resistivity of the surface water in the inverse model; accurate surface water information was inputted into our models with the possible exception of frozen periods where river ice could have altered the geometry of the surface water layer.

Line 486: What about a synthetic model example?

Response: At this point we will limit our results and discussion to the field measurements. A synthetic study would be very informative and could be considered in future work.

Line 510: How does geoelectrical transience translate into hydrological processes?

Response: Changes in electrical resistivity of an area over time indicate variations in the electrical properties of the pore water (the only dynamic component of the system). Changes can occur as a result of temperature, specific conductance, or saturation. Surface water and groundwater typically exhibit distinct electrical properties (Figure 8). Whether these properties can be exploited in bedrock environment using surface geophysical methods has not previously been explored. Our study aims to assess the utility of surface geophysics, while also examining the utility of temperature and EC fluctuations to infer hydrological processes (e.g., thermal conduction, groundwater-surface water exchange, and fracture connectivity) in a bedrock river of varying morphologic conditions.

Line 511 – 452: The conclusions section contains substantial summary and could be reworked for improved focus.

Response: We agree with this assessment. We will rework the conclusions to better highlight the specific contributions of this study with respect to the method and application in bedrock river environments, as well as elements of the conceptual model described in the introduction.

Figure 5: The purpose of this figure is unclear and I suggest that it could be deleted. The A/B/C/D locations are already indicated on Figure 12; the river stage information is presented on Figure 4; the location of the model block midpoints does not appear to be substantially important to the manuscript.

Response: We agree that Figure 5 is not necessary.

Figure 9: Perhaps showing only the difference between these two maps would make interpretation easier? If not only difference, then perhaps just adding a third difference panel. Also, isocontour labels are too small to read.

Response: This figure can be modified to also show the change or difference between data sets. The labels on the contours can be increased in size.

Figure 10: What is the model error relative to the measurement errors? What is the purpose of showing these vast bulk averages when that eliminates any of the valuable spatial information yielded by using tomographic methods? Figure 12 seems to be much more useful than this.

Response: Quantification of measurement error is not straightforward. A precise measurement is not the same thing as an accurate measurement. It is generally easy to obtain a precise measurement by stacking. Imprecise measurements are negated as outliers. However, the accuracy of the measurement has more to do with site conditions (e.g., heterogeneities, anisotropy). Our use of a Wenner array (selected for its higher signal quality) limited our ability to assess the accuracy of the measurements. Other arrays, like the dipole-dipole, permit the collection of reciprocal data, which can be used to quantify measurement error.

The primary purpose of this figure was to the present the time series in its entirety for each location, including the min/max/median value, model RMS error, data removed and relative position of the selected resistivity snapshot shown in Figure 11. The figure also summarizes the field conditions for each measurement (e.g., unfrozen, partially frozen, frozen) which is important for the interpretation of data in Figure 11 and Figure 12. Figure 12 does not provide this information.

Figure 11: Very nice layout and presentation of this figure, however certainly this needs to be replotted to show the difference between (b) through (h) relative to (a) in both columns

Response: This comment is understandable; however; we did not apply a time-lapse inversion due to the large temporal sampling interval and varying data quality (removed data points) over the course of the study, both of which degrade value of time-lapse inversions. In our opinion, presenting the data as absolute resistivity f (rather than relative % change in resistivity) is more representative of the site conditions. At this point we would prefer to maintain the current layout. Figure 12 provides an indication of the relative change in resistivity for specific region in the model at each site.

References:

McNeill, J. D.: Electromagnetic terrain conductivity measurement at low induction numbers, Geonics Ltd. Technical Note, TN-6, 1980.

Van Dam, R.L., B.P. Eustice, D.W. Hyndman, W.W. Wood and C.T. Simmons: Electrical imaging and fluid modeling of convective fingering in a shallow water-table aquifer, Water Resources Research, 50, doi: 10.1002/2013WR013673.