We appreciate the review by referee Prof. Haitjema who stated that the paper is generally well written and offers a useful case study of hydrology of interest to readers concerned with the hydrology of the Konza LTER and similar ecosystems. A number of comments regard the validity of the calibrated stream resistances. These comments are discussed here and changes to the manuscript are put forth to further clarify its contributions.

Comment #1: In fact, I am a bit puzzled as to the methodology of the authors regarding the assessment of these streambed resistances.

Response #1: It is our opinion that the rationale for the choice of our modeling approach needs to be better explained both here and in the manuscript.

The following are unknown parameters that are to be calibrated through model interpretation:

- Hydraulic conductivity of the three zones identified in figure 3
- Resistance of streambeds

The values that we have to use in model calibration are:

- Observed groundwater elevations in figure 6
- The net discharge from surface water to groundwater through the ephemeral streambeds of the Kings Creek watershed in equation (5) and found to be on the order of 100s of m^3/day (line 501) from conservation of mass of surface waters
- The net discharge of groundwater to surface water as baseflow in perennial streambeds of Kings Creek watershed in equation (6) and found to be on the order of 50m^3/day (line 534) from field observation.

A logical method for model calibration might seem to apply PEST to choose the best values of hydraulic conductivity in the three zones and also the value of resistance in streambeds to match the observation wells. This would have given a single value for the resistance of the streambed. This value of resistance might or might not have satisfied the observed discharges. In fact, we wouldn't even know which portions of the stream had water going down and which had water coming up until we had a MODFLOW solution everywhere to get the depth to water. Only then could we integrate the ephemeral portions (by summing the cells with water going down) to see if its discharge matched line 501 and the perennial sections (with water coming up) to see if its discharge matched line 534.

Instead, what we decided to do was to specify the value of resistance to a set of representative values and then solve the model at each fixed value by letting PEST find the best values of hydraulic conductivity to match observation heads. For each of these simulations we then integrated enhanced recharge and baseflow, graphed the observed groundwater/surface water fluxes in figure 9, and picked off the values of resistance that best matched observed discharges.

To our knowledge, this is the first time this approach has been used to establish resistance of streambeds. This method has added advantage of a graphic representation of the sensitivity of discharges to streambed resistances.

Proposed change to manuscript: This needs to be clarified in the manuscript. We propose to change the abstract (line 17) from

"Simple log-log relations correlate the enhanced recharge beneath ephemeral upland streams and baseflow in perennial lowland streams to the unknown resistance of the streambeds."

To

The resistance was set to fixed representative values during model calibration of hydraulic conductivity, and simple log-log relations correlate the enhanced recharge beneath ephemeral upland streams and baseflow in perennial lowland streams to the unknown resistance of the streambeds."

Additionally, the following will be added at line 430:

“The actual value of resistance at Konza will be chosen later by matching the enhanced recharge (5) in the ephemeral streams and the baseflow (6) in the perennial streams.”

We will also add the following paragraph at the end of the discussion section on line 564:

“The model calibration was performed to determine the:

- Hydraulic conductivity of the three zones identified in figure 3, and
- Resistance of streambeds

such the model reproduced the:

- Observed groundwater elevations in figure 6,
- Net discharge from surface water to groundwater through the ephemeral streambeds of the Kings Creek watershed in equation (5) and found to be on the order of 100s of m^3/day from conservation of mass of surface waters, and
- Net discharge of groundwater to surface water as baseflow in perennial streambeds of Kings Creek watershed in equation (6) and found to be on the order of 50m^3/day from field observation.

Our calibration method specifies the value of resistance to a set of representative values and then solves the model at each fixed value by letting PEST find the best values of hydraulic conductivity to match observation heads in table 4. For each of these simulations we then integrated enhanced recharge and baseflow, graphed the observed groundwater/surface water fluxes in figure 9, and picked off the values of resistance that best matched observed discharges. To our knowledge, this is the first time this approach has been used to establish resistance of streambeds.”

<u>Comment #2</u>: Unless I missed this in the paper, model calibration has been limited to comparisons with groundwater elevations in a set of monitoring wells and perhaps base flow at one stream gage at the outlet of the Konza LTER. Alternatively, calibration might have involved a set of stream gages distributed over the model domain, but those are not available I surmise.

<u>Response #2</u>: Actually, we developed estimates of the mean annual recharge occurring across the region using the EPIC model for our study. Thus, we know how much water is going into the aquifer system through recharge, the primary source of groundwater. What is not known (before this manuscript) is how the streams are influencing the water budget: where does discharge occur, where does enhanced recharge through streambeds occur, and what are the controls (measured in terms of streambed resistance)? If we know the recharge and we know the heads then we should be able to develop accurate estimates of hydraulic conductivity since head can be calibrated to within a ratio of recharge over hydraulic conductivity, and we know the recharge. This is substantiated by Haitjema (1995):

“Rule 3.12 Trying to match modeled groundwater mounding with observed groundwater mounding provides insight in the ratio of recharge over transmissivity (N/T). By itself, however, it cannot lead to an estimate of either one of these parameters individually.”

The stream gauging station on Konza was not used during model calibration by PEST, but for interpretation of the water budget at Konza.

<u>Proposed change to manuscript</u>: We propose to clarify this by adding the following to the introduction at line 100:

“Estimates of mean annual recharge occurring across the region will be developed, as well as understanding of how the streams are influencing the water budget.”

And also to modify the sentence on line 173 from

“The purpose of this study is not to quantify the seasonal and interannual forcings of bank storage and hyporheic flow, but to study long-term sustained fluxes through groundwater and baseflow to streams.”

To

“The purpose of this study is not to quantify the seasonal and interannual forcings of bank storage and hyporheic flow, but to study long-term sustained fluxes through groundwater and baseflow to streams, and how the properties of streambeds control groundwater/surface water interactions.”

Comment #3: The authors point out that stream recharge (“enhanced recharge” in their paper) is very limited, which I take to mean that it also has a limited impact on groundwater elevations. Under these circumstances it seems difficult to calibrate for stream resistances, particularly when trying to distinguish between resistances of 10,000 days and 100,000 days. In both of these cases the streams have virtually no impact on the groundwater flow regime, even though their (very small) recharge to the groundwater differs by an order of magnitude, of course. This is confirmed in table 4, where the average differences in head errors (at monitoring wells) for these two resistance values is only 1 cm, not a very discriminating result. Also note, in the same table, that the average difference in error between c=100 days and c=1,000 days is also very small: 4 cm.

Response #3: Yes, we agree. The model results show that the resistance to flow required to match observed discharges of stream flow result in a limited impact on overall water balance and groundwater elevations. We believe that this case can be made stronger in the paper by discussing the average differences in error between modeling scenarios as pointed out by this referee.

Proposed change to manuscript: The following will be added at line 436:

“Note that the model results indicate that only small changes in conductivity and RMSE in table 4 occur when the streambed resistance changes from 100,000/day to 10,000/day. As the streambed resistance takes on values of 1,000/day to 100/day the surface water interactions through streambeds begin to noticeably change calibrated values of hydraulic conductivity.”

Comment #4: These results were obtained by not only changing the stream resistances, but simultaneously also changing the hydraulic conductivities in each of the three zones in which they differ. I fear that the latter variations may have been more important (more impact) than the stream resistance values

Response #4: The model results in table 4 partially substantiate this. There are only small changes in hydraulic conductivity when the streambed resistance changes from 100,000/day to 10,000/day. As the streambed resistance takes on values of 1,000/day to 100/day the surface water interactions through streambeds begin to noticeably change calibrated values of hydraulic conductivity.

Proposed change to manuscript: We believe that this has been addressed in proposed changes to previous comments (particularly changes for comment #1 and #3).

Comment #5: The authors did compare the use of different streambed resistances with simulated base flow for the Konza LTER (Fig. 9) at one gage. Since the base flow changed it implies that the groundwater divide moved in response to varying stream resistances (stream recharge). Specifically, a larger stream resistance increases the groundwater elevations and with it the contributing watershed size. This could, in fact, offer a more sensitive calibration mechanism, but only if all stream resistances are changed in tandem (as they are – in fact all were assumed the same, I believe).

Response #5: Yes, we agree. Unfortunately, we do not have data to substantiate to location of the groundwater divide. Perhaps future research and field observations can help to identify this.

Proposed change to manuscript: No changes requested by the referee.

Comment #6: In summary, I am a little suspicious as to the meaning of the stream resistance assessment in this modeling exercise. The assumption that all stream resistances are the same seems one of convenience rather than reality and may defeat the accuracy attempted with the automated calibration process (PEST). It would have been interesting to keep all hydraulic conductivities the same and then vary the stream resistances, just to see how they, *by themselves*, affect the modeling results.

Response #6: We believe that we have addressed this concern in the response to comment #1. Furthermore, the studies we could find that used GFLOW to study streambed resistance modified both hydraulic conductivity and resistance values,

- Krohelski, Rose, and Hunt (2002) USGS WRI 02–4034: "The GFLOW model was calibrated by trial-and-error; that is, by varying hydraulic conductivity and stream resistance until there was a reasonable match between measured and simulated ground-water levels (heads) and streamflows."
- Hunt, Juckem, and Dunning (2010) USGS SIR 2010–5049 used GFLOW and PEST to calibrate regional recharge value, the local hydraulic-conductivity zone near the Hayward Airport, and the local streambed resistance. The emphasis of the calibration was not on simulating the local groundwater levels or the distribution of groundwater flow within the Northern Tributary, but to simulate the distribution of flow between the entire Northern Tributary (as represented by the most downstream base-flow measurement) and the other potential sink for water recharged in the Hayward Airport (note that GFLOW model parameters automatically give the contribution to stream flow from groundwater along a stream reach, whereas MODFLOW does not).

This is a very astute observation. Ultimately, whether we vary hydraulic conductivity for fixed values of resistance (as we did in the paper) or vary resistance for fixed value of hydraulic conductivity (as proposed in this comment), we should get the same results when we match the conditions of discharges.

Proposed change to manuscript: We believe that the above changes to the manuscript address this concern.